

DOCUMENT RESUME

ED 039 517

AL 002 411

AUTHOR Fasold, Ralph William August
 TITLE Noun Compounding in Thai.
 PUB DATE 69
 NOTE 226p.; Ph.D. Dissertation, University of Chicago, 1968
 AVAILABLE FROM A limited number of copies available from the author, Sociolinguistics Program, Center for Applied Linguistics, 1717 Massachusetts Ave., Washington, D.C. 20036
 EDRS PRICE EDRS Price MF-\$1.00 HC Not Available from EDRS.
 DESCRIPTORS Deep Structure, Linguistic Theory, *Nominals, Phrase Structure, Surface Structure, *Thai, *Transformation Generative Grammar, Transformations (Language), *Transformation Theory (Language)

ABSTRACT

The present study, a slightly revised version of the author's 1968 Ph.D. thesis presented to the University of Chicago, investigates compound formation in Thai. Chapter 1 summarizes the transformational generative theory on which the study is based, discusses the concept that Thai is a "simple" language in comparison with English, and briefly outlines the structure of Thai noun phrases. Chapter 2 outlines the development of transformational grammar and presents the base component of a transformational grammar. Chapter 3 discusses the use of transformations and describes the transformational rules needed to generate the structures directly underlying the compounds themselves. Chapter 4, working from the assumption that most Thai noun compounds are derived from noun phrases in which sentences have been embedded as relative clauses, demonstrates how certain structures generated by the grammar proposed in the earlier chapters can be converted into Thai noun compounds. Compounds are analyzed by type. Chapter 5 is dedicated to the small subset of noun compounds considered to be learned compounds and offers solutions to the problems presented by such forms. [Document not available in hard copy because of marginal legibility of original.] (FWB)

ED039517

NOUN COMPOUNDING IN THAI

BY

RALPH WILLIAM AUGUST FASOLD

ARLINGTON, VIRGINIA

1969

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE
PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION
POSITION OR POLICY.

AL 002 411

TO GAE

Who Shares Everything

Preface

This manuscript is essentially a reprinting of my doctoral dissertation submitted to the University of Chicago (1968). This version is abridged from the dissertation in two ways. The appendix to the dissertation, on reduplication, has been eliminated and is in preparation for separate publication. The long lists of examples of compound types which appeared in the dissertation have been limited to ten examples of each type (for a few types there are fewer than ten examples).

The development of the theory of generative grammar is notoriously rapid. Publications dealing with the theory or with its application to real languages are often outdated by the time they are released. The discussion of the development of generative theory in chapter 2 terminates at a particularly transient point. The reader would be ill-advised to take chapter 2 as an up-to-date treatment of the theory. On the other hand, the stage in the development of the theory of deep structure now espoused by McCawley, Lakoff, Langendoen, and others, which is discussed in chapter 2, may be of interest from a historical point of view.

This study of noun compounding may well raise doubts as to the correctness of deriving all noun compounds in natural languages from full sentences via relative clauses. In a number of instances, it is impossible to get informants to agree on what the main verb or other elements of the underlying sentence should be. Perhaps at least some compounds should

be directly derived from semantic representations.

In particular, the plant-name compounds at the end of chapter 4, and the learned compounds in chapter 5 are not very satisfactorily treated. The learned compounds in chapter 5 are particularly interesting since a very similar problem exists in English with regard to certain words of Greek and Latin origin. To my knowledge, no successful analysis of these English words has been proposed.

In undertaking this study, I am indebted especially to Professor Eric P. Hamp of the University of Chicago, who encouraged me to work in Thai syntax and who advised me throughout the preparation of the dissertation. Special appreciation is also due to Professor James D. McCawley in whose stimulating classes I gained an understanding of generative transformational linguistics and who also made a number of helpful comments on the dissertation. I also wish to thank Professor William J. Gedney of the University of Michigan for graciously taking the time to discuss several matters in connection with my work and also for making available to me Miss Nisa Udomphol's master's thesis.

Without the help of several Thai nationals who willingly served as informants, this study would have been totally impossible. Miss Suphis Thanissom was of immense help in the early stages in familiarizing me with Thai structure. Mr. Montri Chenwidthayakan spent many long hours with me dictating anecdotes and other material on which much of the analysis

is based. Mr. Wichit Sirisamphan and Mr. Kasian Chongsarit, who are mentioned several times in chapter 4, gave invaluable help in supplying judgments of grammatical acceptability and in suggesting derivational sources for some of the compounds. Special thanks is due to Miss Nisa Udomphol, a linguist in her own right, who nevertheless condescended to act as my informant during her time as a student at Chicago. It will be obvious in reading chapter 4 that I am in her debt both for her summaries of native Thai grammars and for some of her own insights into the structure of Thai compounds, to be found in her master's thesis.

I also wish to thank Mrs. Myrtle Tozar who typed the final draft of the dissertation with sufficient speed and accuracy to enable me to meet a very rapidly approaching deadline.

Perhaps no one deserves more appreciation than my wife, Gae. Her hard work, constant encouragement, and willingness to do without made my whole graduate career possible. In spite of the fact that she is unfamiliar with both Thai and the symbol conventions of linguistic description, she typed the entire second draft of the dissertation from my all-but-unreadable rough draft. The quality of her work was so excellent that only minor editorial changes were necessary in the final draft to satisfy the very demanding University of Chicago Dissertation Office. As if this were not enough, she heroically typed virtually the entire dissertation again, this time on mimeograph stencils for the present version. It is to Gae that I dedicate this manuscript.

RWF
July 5, 1969

TABLE OF CONTENTS

	Page
LIST OF ILLUSTRATIONS	vii
Chapter	
1. INTRODUCTION	1
Introduction	
Summary of the Theory	
Thai: a "Simple" Language	
Noun Phrase Structure	
2. INPUT TO TRANSFORMATIONS	26
Introduction	
<u>Syntactic Structures</u>	
Katz and Postal	
<u>Aspects of the Theory of Syntax</u>	
"The Base Component of a Transformational Grammar"	
3. TRANSFORMATIONS	52
Introduction	
Justification for Transformations	
Conventions	
The Rules	
4. ANALYSIS OF COMPOUNDS	86
Introduction	
Other Work on Thai Compounds	
Methodology and Data Sources	
Analysis of Compounds by Type	
5. LEARNED COMPOUNDS	185
Introduction	
Possible Solutions	
Informant Behaviour	
Tentative Solution	
Examples	
BIBLIOGRAPHY	217

LIST OF ILLUSTRATIONS

Figure	Page
1. Structure of the English Sentence: "Henry is hungry."	14
2. Structure of: "John is running."	16
3. Structure of: " <u>cha</u> · <u>kamlan</u> <u>wín</u> "	17
4. Structure of: "John has run."	19
5. Structure of: " <u>cha</u> · <u>wín</u> <u>lé·w</u> "	20
6. Operation of the Amvivalent Verb Transformation	60
7. Operation of Relative Pronoun Insert Transformation	62
8. Operation of the Redundant Noun Deletion Transformation	64
9. Operation of the Relative Pronoun Pre-position Transformation	65
10. Operation of the Indirect Object Transformation	67
11. Operation of the Adverbial Promotion Transformation	71
12. Operation of the Possessive Transformation	76
13. Operation of the Nominalization Transformation	79

Chapter 1

Introduction

1.0 Introduction. This study is addressed to the problem of investigating compound formation in Thai. Various types of noun compounds will be treated, including a subset of noun compounds which we call "subcompounds" in which one or more members are bound forms, and reduplication as a special case of noun and verb compounding with phonological implications. In addition, we shall deal with learned compounds of Indic origin. This rather substantial area of Thai grammar will be described in terms of transformational grammar. It soon becomes apparent that a view of language that goes substantially beyond simple description of observational phenomena is necessary to give a unified account of the various features we shall call compounding. If the relationship between compounds and other constructions in the language is acknowledged, it develops that less than a dozen rules will account for all the kinds of compounding, and that all these rules are rather low-level.¹ Put in other words, Thai compounds are only a short step removed from other, full sentence constructions.

2.0 Summary of the Theory. The investigation will be in terms of transformational grammar as initiated by Noam Chomsky.² As presented in Aspects of the Theory of Syntax,

grammar is a tripartite structure, consisting of semantic, syntactic, and phonological components. It will be the syntactic component which will be of interest in this study, so we will have little to say about the semantic and phonological components.³

The syntactic component is itself divided into two sub-components. One has as its purpose the specification of the form sentences take at the deepest level. This structure is very different from that which the sentences will have on the surface. In Chapter 2 we will discuss the development of this part of the syntactic component in detail and present the deep forms of the types of Thai sentences necessary for our analysis of noun compounds. The other component has as its function the restructuring of these forms into structures which immediately underlie spoken sentences. In the conception of generative grammar used here, we assume that the first sub-component presents to the second a rather small set of structures which may conveniently be represented as inverted tree diagrams. The points at which the branches connect to each other are called nodes and each node bears a label which designates the grammatical structure of that part of the tree which branches below it and is connected to it. The transformations in the second subcomponent operate by adding, removing, or deleting sections of these trees. More details concerning the nature and operation of transformations will be given in Chapter 3.

3.0 Thai: A "Simple" Language. To a European grammarian, especially one of a traditionalist-prescriptivist bent, Thai seems to be a very simple language. Gordon H. Allison, author of a Thai grammar of this type, makes the observation: "Thai grammar is easier in many respects than that of European languages".⁴ Another such grammarian, Peter A. Lanyon-Orgill, goes so far as to say:

Thai, in common with the other languages of the Far East, does not possess any grammar in the European sense of the term... Consequently, the study of Thai grammar largely resolves itself into a study of a section of the vocabulary of the language, and much of the grammatical terminology to which Western scholars are accustomed may be forgotten as being irrelevant to our study as so many of these ideas are quite foreign to the Thai mind.⁵

Sometimes, these grammarians view this simplicity as a form of imprecision or disorganization:

Formerly, Thai grammar was rather loose and disorganized; however, in modern times Thai grammarians have established Thai grammatical rules patterned to some extent upon those of European language.⁶

Campbell and Shaweevongse, in discussing what they call the "continuous tense", make the comment:

You will not find the continuous tense used very much in conversation as generally the Thai people are not very much concerned about the finer shades of meaning and see very little difference between "I buy a tomato" and "I am buying a tomato". As long as their meaning gets across they are prepared to take considerable liberties with formal grammar.⁷

A glance through these grammars shows in just what aspects Thai grammar shows itself to be simple:

Thai verbs are never conjugated; variations in person or tense are shown by other words or by inference.⁸

Articles (a, an, the) are not necessary in Thai... Indefinite adjectives (some, any), the same as articles, are often not necessary in Thai... Nouns have only one form. Number, gender, and case are all shown by helping words or by inference.⁹

Adjectives cannot follow a verb in Thai. Therefore, there are no predicate adjective constructions in Thai (such as: "He is good"). A Thai sentence is often complete without a verb, whereas a "be" verb would probably be necessary in the English translation.¹⁰

The general arrangement of a simple Thai sentence is Subject-Verb-Object as in English... There is no inflection of nouns, pronouns or verbs... Inanimate objects have no gender and where it is necessary to indicate gender in relation to animate objects this is usually done by the addition of extra words.¹¹

The "simplicity" of Thai, then, seems to consist of:

- 1) Lack of conjugated verb forms indicating person or tense (number might also be added), 2) Lack of inflected noun forms indicating number, gender, and case, 3) Lack of articles and indefinite determiners, 4) Lack of case forms of pronouns, 5) Lack of copula in Thai translations of English predicate adjective constructions, and 6) Simple (and what is more interesting, stable) basic word order in sentences.

On the other hand, other pedagogical grammars of Thai, written by grammarians who may be described as structuralists, deny that Thai is any simpler than English or any other European language. In the first "Word Study" in her grammar book, Mary Hase states:

The fact that a single word can correspond to so many different things in English may cause you to feel at first that Thai is not as definite and precise as English. But as you progress in your understanding of the language, you will notice that whenever it is necessary to be precise about number or tense, Thai can be precise, but that when such precision is unnecessary Thai is not bound, as English is, to be precise about it.¹²

But it is pointed out to the student that Thai can make precise distinctions that English does not make:

In the preceding paragraph we noted that English insists on making distinctions in number and tense which are not compulsory in Thai. In its turn we find that Thai makes certain distinctions of other types which are not compulsory in English. One of these distinctions is that in Thai certain words are used by men while others, of corresponding meaning, are used by women.¹³

The argument seems to be that Thai has all the grammatical machinery necessary to express anything that English can, but does not express some things in every sentence which must always be expressed in English. Furthermore, Thai has a regular system for pronouns which designates the sex of the speaker which must always be utilized where English lacks such a system. The implication seems to be that on balance both languages are about equally complex. In his grammar book, Edward Anthony states the same argument more succinctly:

From the lack of verbal endings and the identity of singular and plural nouns, you may be led to believe that Thai has no grammar, or a very rudimentary one. Nothing could be further from the truth. Thai has as many and as complicated grammatical subtleties as does English, but signals them in a different way.¹⁴

From this point of view, Thai is not easy at all; it just has different ways of expressing things and has grammatical

features of interesting and complex types which English and European languages lack. It is our contention that both views have merit. It appears that there are two factors at work in this issue: 1) from the presumably universal inventory of syntactic elements, Thai and English select partially overlapping and partially different subsets, and 2) there are a number of places at which English underlying structures will look strikingly like Thai surface structures, but its surface structures will be very different. It is this second factor which will prove most interesting and will provide a deep explanation for the "simple" impression Thai makes on the speaker of a European language. We will attempt to show that one way in which Thai is simpler than English is in that it lacks certain transformations which make English surface structures very different from their corresponding deeply underlying forms. The first factor is the one which stands behind the structural grammarian's statements. First we shall see in detail how it applies to Thai grammar.

3.1 Lack of Grammatical Elements.

3.1.1 Lack of Morphemes. Thai can be said to lack certain morphemes which are present in languages presumed to be more complex. An obvious case is the one involving articles.¹⁵ Articles are "not necessary" in Thai simply because Thai has no such elements in its inventory of morphemes. A somewhat less obvious case involves tense. In transformational grammar, it has been found that the most economical statements involve

treating tenses as abstract entities in their own right, although tenses are often expressed phonologically as affixes on verbs. In An Integrated Theory of Linguistic Descriptions, for example, Katz and Postal present a tree structure for the English sentence "John sleeps." in which the terminal symbols are John-Present-sleep, indicating that at this level of analysis, Present has the same sort of status as does John and sleep.¹⁶ Since in Thai time of action cannot be indicated except by the use of time adverbs (although it may be implied by the use of devices discussed below), it is evident that Thai simply lacks elements like Present altogether.

It is true, as Haas and Anthony would argue, that Thai has classes of morphemes comparable to articles and tense morphemes which English lacks. One interesting class of this type, which we shall not have occasion to discuss in detail, is a class of sentence particles. For instance, there is one type which is added to a great many sentences to indicate the speaker's status with relation to the hearer. These are extremely hard to define or translate, but we will give some of Richard Noss' examples, along with his definitions and illustrative sentences:¹⁷

- | | |
|-------------|--|
| <u>kháp</u> | 'Male speaking to superior, elder, or non-intimate equal person' |
| <u>khâ</u> | 'Female speaking to superior, elder, or non-intimate equal person' |
| <u>câ</u> | 'Person speaking to inferior or younger person' |
| <u>wâ</u> | 'Person speaking rudely or to intimate equal' |

In the examples below, I have identified the particle by the use of the abbreviation ptl. in the interlinear translation (which I supply; Noss gives only the smooth translation) and have not tried to translate them. Noss does not give any examples of sentences with câ. Speakers are identified by letters of the alphabet. I have translated Noss' transcription into the one I shall use throughout this study.

B. sây sya dichán pay sí khâ
 put-on blouse I go why-don't-you ptl.

"Wear my blouse, will you?"

C. phôm hă mây phóp khráp
 I look not find ptl.

"I can't find it."

C. khray than thûaykê·w tê·k wâ
 who make drinking-glass be-broken ptl.

"Who broke the glass?"

3.1.2 Lack of Syntactic Features. In the conception of generative grammar with which we are dealing, items entered in the lexicon have sets of syntactic features associated with them. Some of these features serve to classify items in certain ways. For example, the English verb 'know' will bear the feature [+activity] signifying that it is an activity verb. The English temporary aspect marker must be marked in its set of syntactic markers so as to keep it from being associated with activity verbs, since there is no sentence *"He is knowing all the answers." Other features specify the contexts in which the item may appear in sentences. To express

the restriction about activity verbs, for example, the temporary aspect marker would require a feature something like $[- \text{---} [+V \text{---} [+activity]]]$ meaning that it cannot precede activity verbs. In this conception, it would also be possible that two languages could have different sets of such features available. In Chomsky's discussion of matters of gender, case, and number in Aspects, number would be an example of this type. Number would be introduced in the base component, which specifies basic structures, by the rule which develops the category Noun. Accordingly, the rule would be:

$$N \rightarrow [\Delta, \alpha \text{ Number}] (\alpha = + \text{ or } - \text{ for English...})^{18}$$

In this notation, Δ is a dummy symbol which is to be replaced by a lexical item whose syntactic features are such that it will be allowed to appear in the structure in which the Δ appears. As we have said, lexical entries involve sets of features and the one which replaces noun dummy symbols will automatically have either $[+Number]$ (i.e., plural) or $[-Number]$ (i.e., singular) added to its inherent set of syntactic features. Transformations later will read these features and insert the associated phonological material; in English the reflex of $[+Number]$ will usually be $[s]$, $[z]$, or $[əz]$, and of $[-Number]$ $\emptyset$. If this is the correct analysis of number in languages like English, then this is an instance in which Thai lacks a syntactic feature which English has.

It is possible to argue that the absence of articles and

tense morphemes are also cases of the lack of syntactic features. If Thai lacks features present in English like [+Article] and [+Tense], it follows that it will lack lexical entries which must be marked with such features.

3.2 Lack of Transformations. More far-reaching and interesting than the comparison of inventories of grammatical elements is a comparison of the transformations between Thai and the well-known European languages. It will be noticed that Thai was said to lack, among other things, tense, number, gender, and case affixes. While all these may seem to be of the same type, only tense and number were treated as an example of a lack of a morpheme (or a feature). The others are most properly viewed as cases of absences of transformations of certain kinds possessed by European languages. As Chomsky points out, case is not relevant to deep structure at all, but serves to mark certain relationships in surface structure. For this reason, it is introduced by transformations.¹⁹ To be sure, if the transformations do not exist, the features which they would introduce will also fail to exist, but it is apparent that the lack of the transformations is fundamental. Obviously, the grammatical relations Subject-of, Object-of, etc. can be found in Thai; it is only the case system for marking them with features like [+Nominative], [+Accusative], etc. which is missing.²⁰

Both case features and the transformations for introducing them are absent in Thai grammar, but in the case of

gender and number, the features are present, but the transformations which associate discrete phonological affixes to them are not. It is clear that Thai has pronouns which must be marked as Second Person, First Person, and Third Person. Similarly, Thai has nouns which must be marked with gender features, especially in the kinship system. The words cha·y and să·w mean 'male human' and 'female human', respectively. When paired in compound formation with phî· 'older sibling' and nó·n 'younger sibling', we get phî·cha·y 'older brother', phî·să·w 'older sister', nó·ncha·y 'younger brother', and nó·nsă·w 'younger sister'. cha·y and să·w function similarly with other kinship terms as well. In at least one usage, kin terms are used outside the kinship system, apparently especially for their gender features. The compounds phô·khrua (literally 'kitchen father') and mē·khrua (literally, 'kitchen mother') are the words for 'male head cook' and 'female head cook', respectively.²¹ Chomsky suggests that features such as person, number, and gender enter into agreement relations in language like German via rules of the form:²²

$$\text{Article} \rightarrow \begin{bmatrix} \alpha\text{Gender} \\ \beta\text{Number} \\ \gamma\text{Case} \end{bmatrix} / \text{---} \dots \begin{bmatrix} +N \\ \alpha\text{Gender} \\ \beta\text{Number} \\ \gamma\text{Case} \end{bmatrix}$$

In this notation, α , β , and γ are variables ranging over integers and the various gender, number, and case features are identified by integers. The above rule specifies that an article be assigned the same gender, number, and case features

as the following noun. Later transformations would assign specific phonological material to the various combinations of these features. Thai, however, lacks such rules.

Thai word order is also cited as a factor contributing to its simplicity. What is remarkable is that Thai word order remains stable under a variety of circumstances under which English sentences undergo rearrangement. For example:

sunák khàt phû·chǎ·y
dog bite man

"The dog bit the man."

is a simple declarative sentence. When questioned for a yes-no answer, it becomes:

sunák khàt phû·chǎ·y rǎi·
dog bite man or

"Did the dog bite the man?"

A content question, even questioning the object, preserves the word order:

sunák khàt ?aray
dog bite what

"What did the dog bite?"

Passive sentences, as such, do not exist in Thai. In order to give a passive import to our example sentence, the whole sentence becomes the object of another sentence with the verb thù·k 'to come in contact with':

phû·chǎ·y thù·k sunák khàt
man come-in-contact-with dog bite

"The man was bitten by the dog."

Rather than rearrangement of the basic sentence, what we have

is the addition of a new sentence, with subsequent deletion from the first one. Clearly, Thai grammar lacks the transformations which effect the word order changes in the corresponding English sentences.

Even more striking than anything discussed so far, is the deep similarity between English and Thai with regard to the construction noticed by Allison in which the Thai sentence containing a predicate adjective lacks a copula verb. Structuralist linguists are generally agreed, and rightly so, that adjectives in Thai are simply a special class of verbs. The copula is not needed in Thai any more than it would be needed before any other verb, for instance, the verb no·n 'sleep'. dek no·n "The child sleeps" and dek ya·y "The child is big" are equivalent in structure and the copula is superfluous in either. But exactly the same sort of analysis has recently been proposed for English adjectives, as well.²³ The copula 'be' originates from outside the adjective predication itself, either by a transformation which inserts 'be' into such structures or by certain deletions from more complex structures. Ross proposes the underlying structure in Figure 1.1 for "Henry is hungry".²⁴ Deletion of the second occurrence of 'Henry' by a very general transformation and subsequent adjustments in the tree structure result in the sentence "Henry is hungry". No matter what answer proves right, it is clear that Thai is no simpler than English at the deepest level (where sentences like "Henry hungry" and dek ya·y both exist). Only

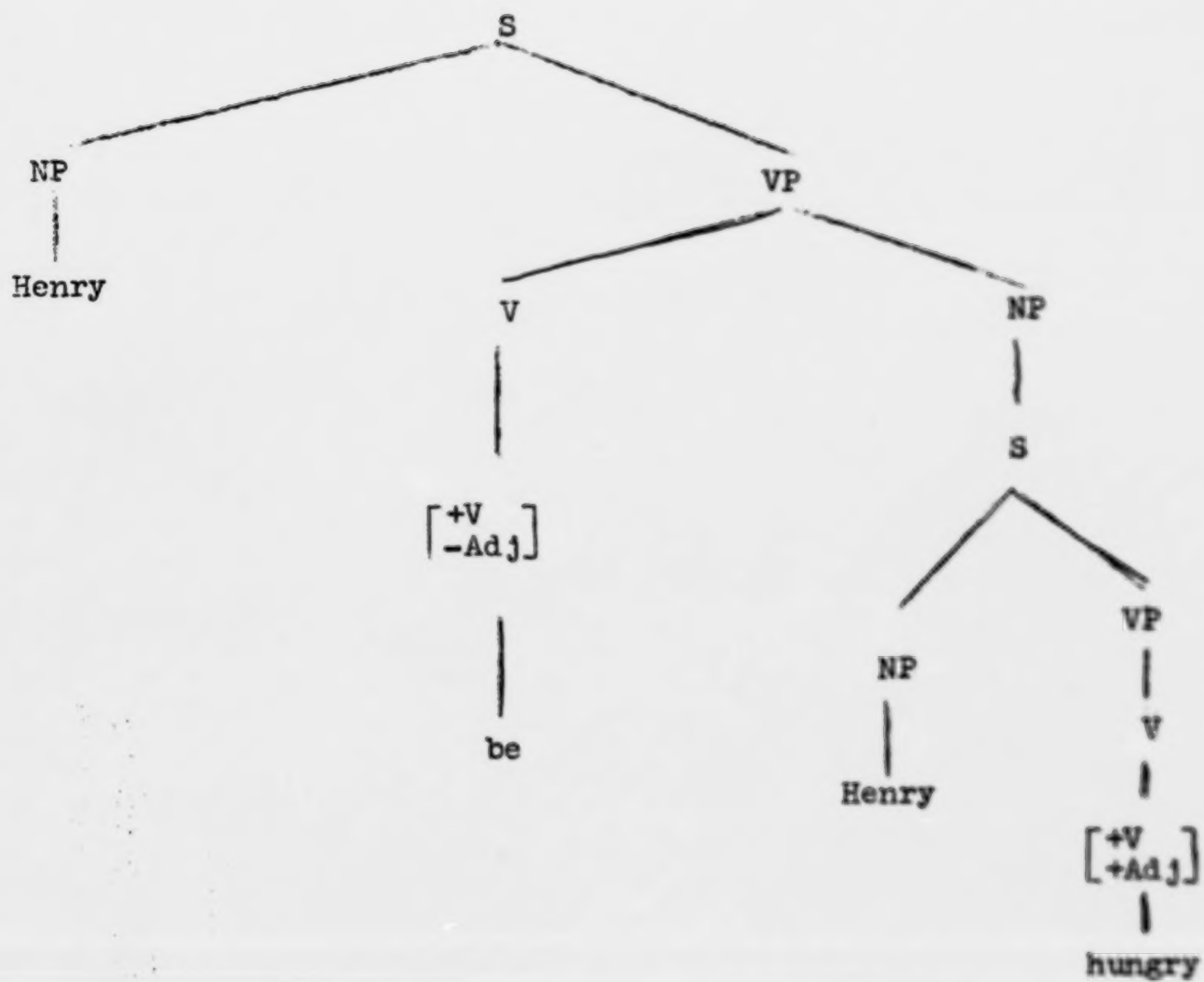
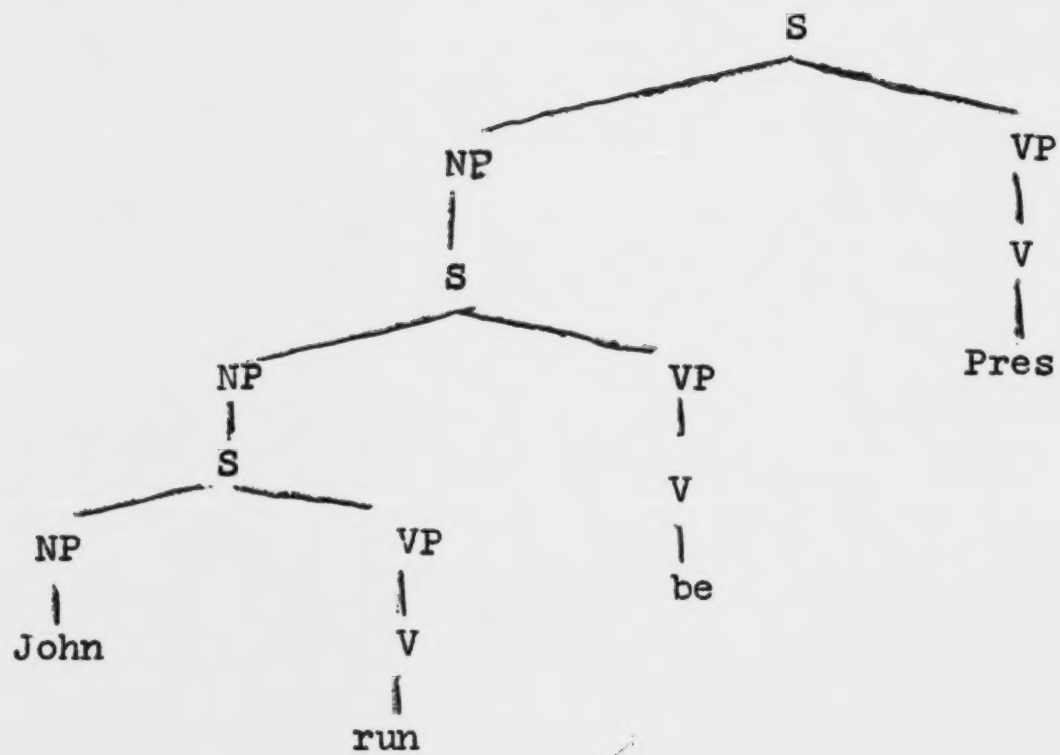


Figure 1. -- Structure of the English sentence: "Henry is hungry."

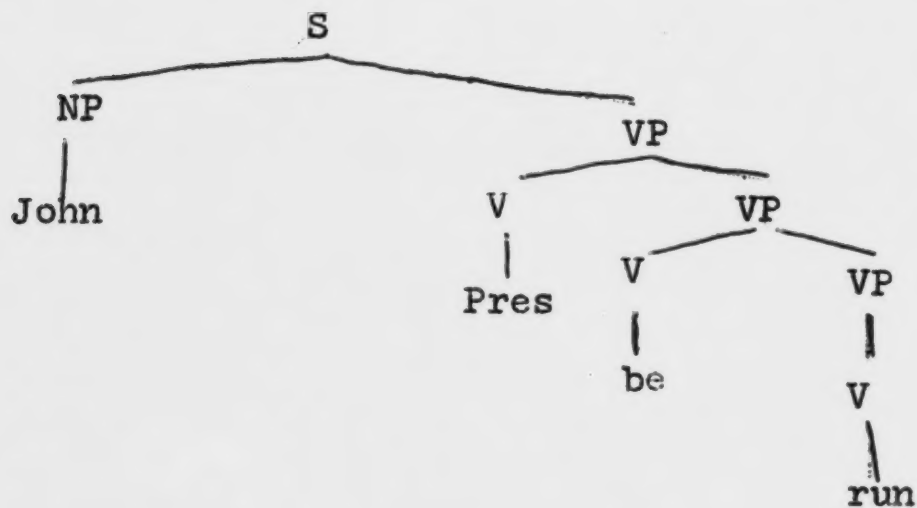
in the transformational development of the English sentences are there complexities unknown to Thai grammar.²⁵

More recent developments in the transformational grammar of English show that another aspect of the simplicity of Thai can be ascribed to transformational developments. In unpublished work, some transformational grammarians have argued that tense and aspect morphemes not be developed from such constituents as Auxiliary, but be treated as verbs which take whole sentences as subjects. In this sort of analysis, the sentence "John is running" would have an early structure like that in Figure 2a. The verbs Pres and 'be' would be subject to a rule of "Verb Phrase Promotion" which detaches such verbs and inserts them in front of the surface main verb. Verb Phrase Promotion would yield the structure in Figure 2b. Later transformations combine 'be' and Pres to form 'is' and insert the suffix 'ing' to 'run'. If the same analysis is applied to Thai, the pre- and post-Verb Phrase Promotion structures (recalling that Thai lacks anything answering to Pres) would be very similar to English, namely those in Figure 3.

So far, so good. The Thai sentence behaves quite analogously to the English sentence, once we have allowed for the absence of tense morphemes in Thai. The difference now is that, except for some possible node relabelings, the Thai sentence is in its surface form. "Cha is running" is pronounced cha· kamlan win; there are no affix insertions or alterations



a.

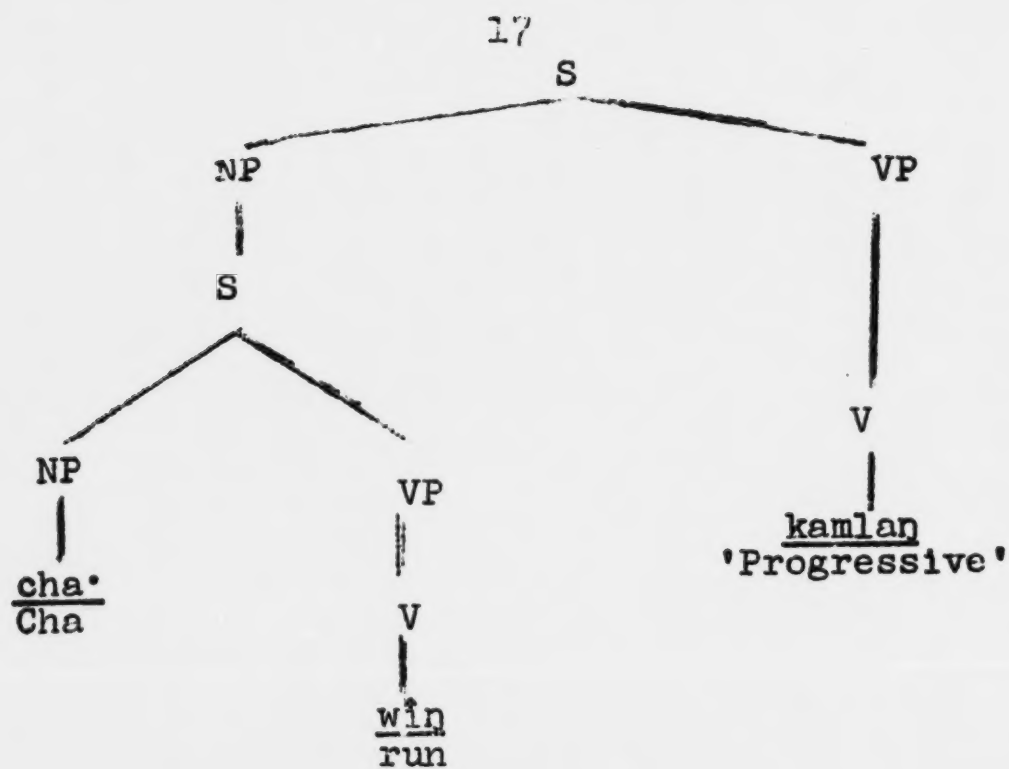


b.

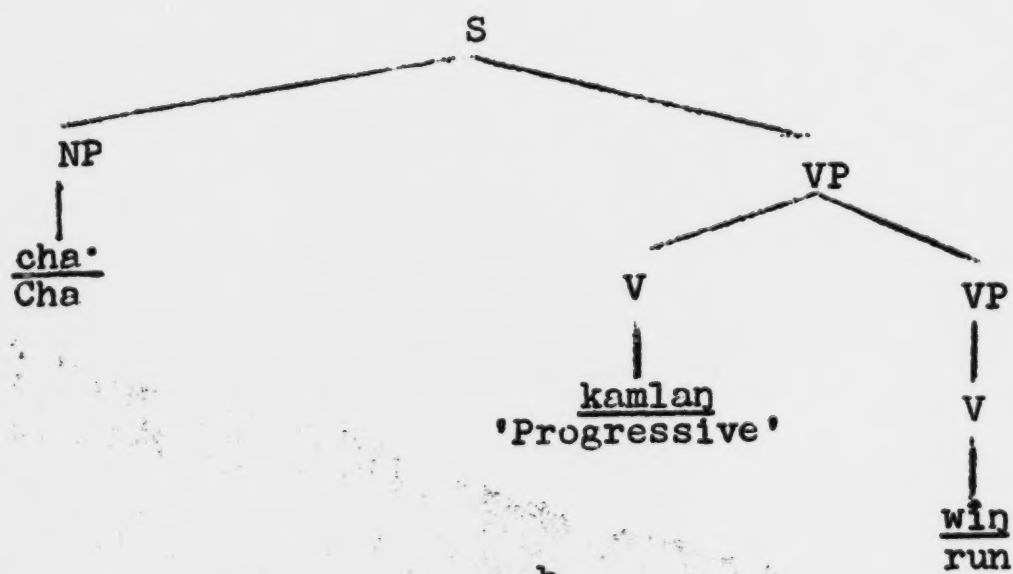
Figure 2. -- Structure of "John is running."

a. Earlier structure

b. Later structure



a.



b.

Figure 3. — Structure of: "cha kamlan win".

a. Earlier structure.

b. Later structure.

of the Progressive morpheme kamlan. Again we see that the simplicity of the Thai sentence lies in its lack of certain rather low-level transformations.

Now let us look at another case. The two structures for "John has run" would be those in Figure 4. The second structure needs some low-level work, such as the spelling of the Completive sentence-verb as 'have' and the combination of 'have' and Pres to form 'has'. In Thai, the deeper structure is that of Figure 5.

On the analogy of the previous sentence, one would expect the Verb Phrase Promotion rule to apply, but perhaps that would be all that is necessary to generate the surface structure. But what actually happens is that the Verb Phrase Promotion rule does not apply and the sentence is now in its correct surface form, except for some node relabeling. The striking result is that the surface structure of the Thai sentence is very similar to the deep structure of the English sentence and its own deep structure, unlike that of the English equivalent. What the Thai deep structure is like is extremely clear, unlike the English case. Notice also that there is no special morpheme marking completive aspect; the regular adjective for 'be-finished' serves in this function.

But what is the source of the difference between the "simple" and "complex" languages here? The answer is that the difference is nearly the most trivial possible. The Thai word lé·w 'be-finished' does not undergo Verb Phrase Promotion, a

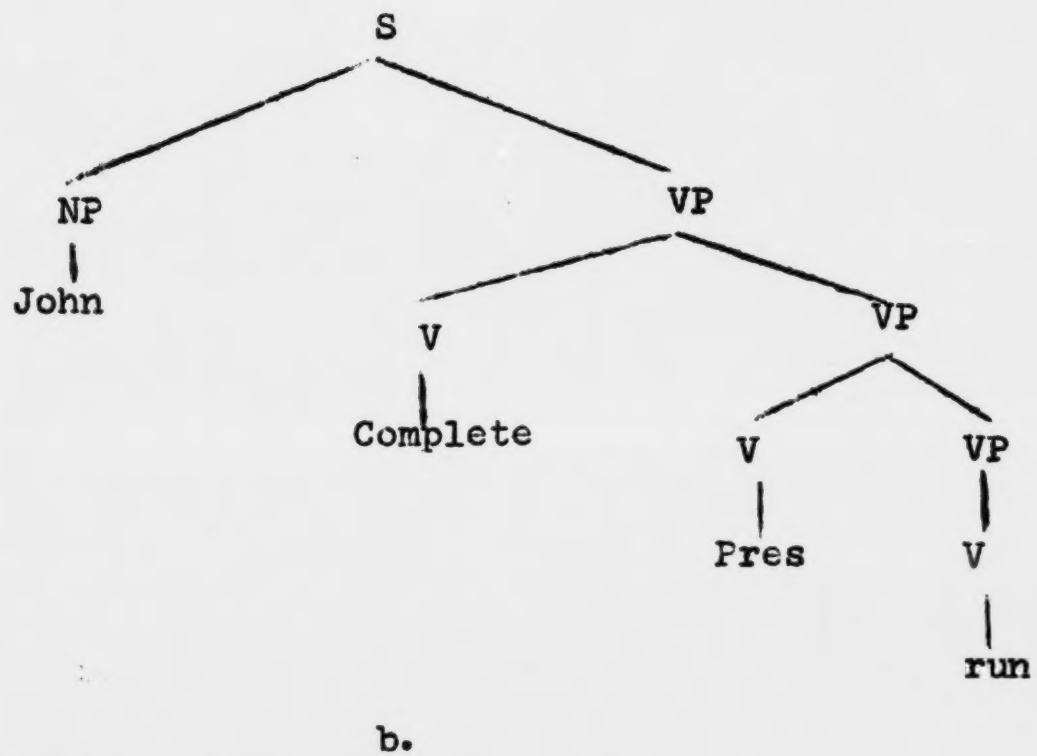
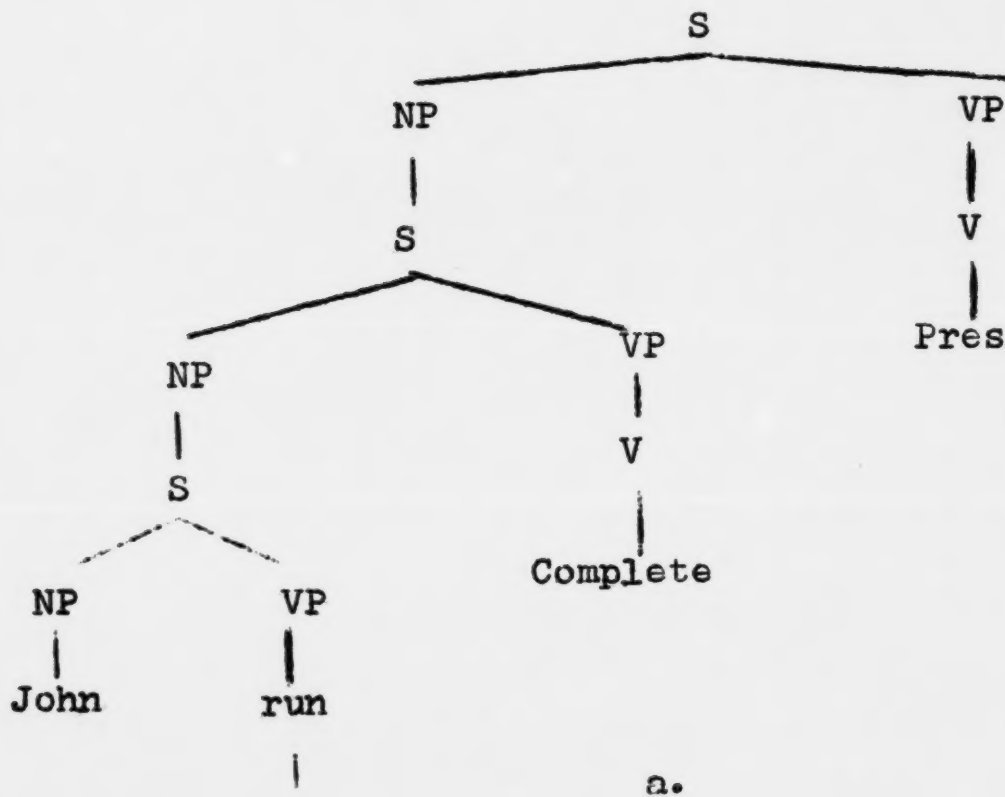


Figure 4. -- Structure of "John has run".

- a. Earlier structure.
b. Later structure.

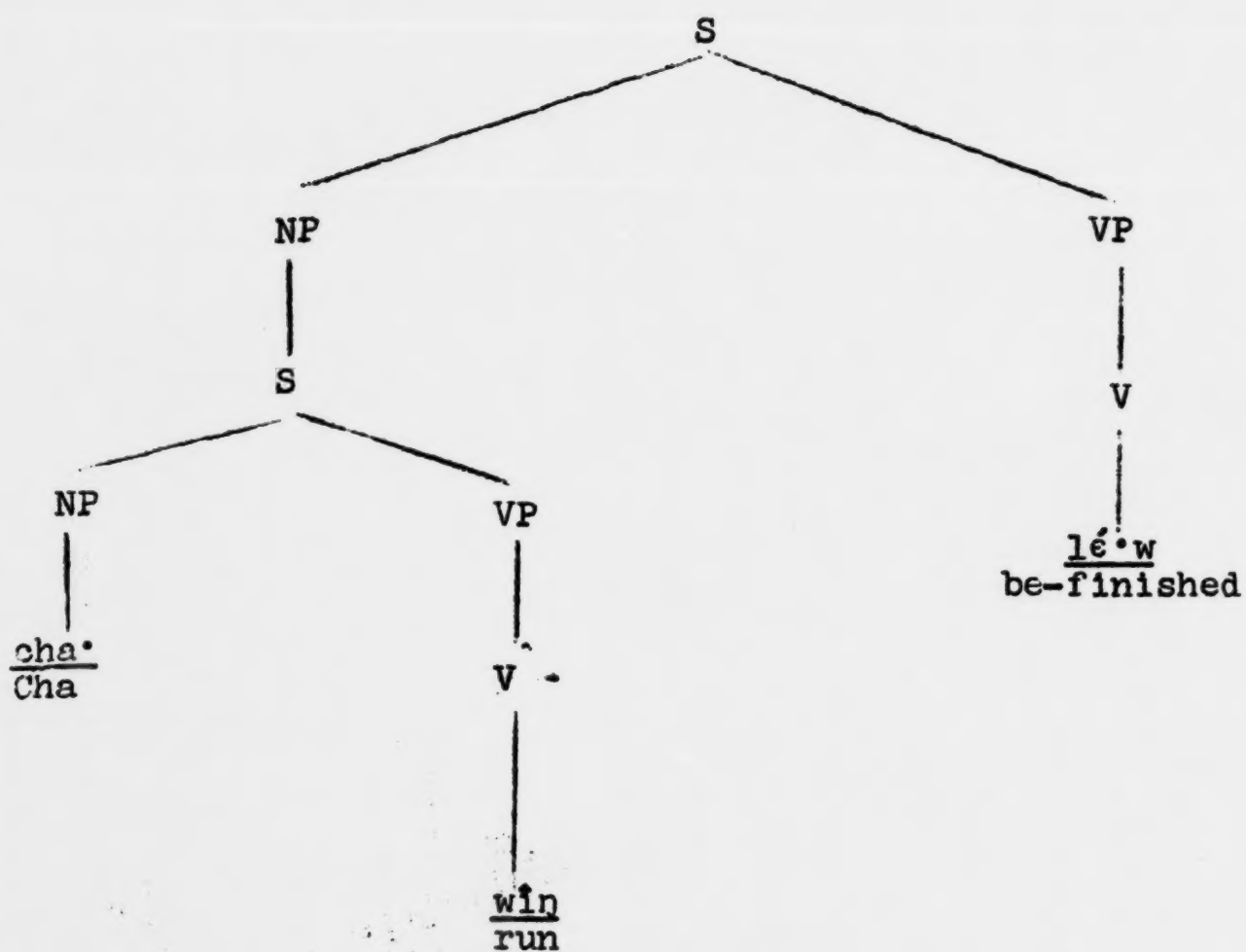


Figure 5. -- Structure of "cha wín lé·w"

rule which both languages clearly have, but its English translation equivalent does.

We have seen that the ultimate source of the apparent simplicity of Thai lies in 1) the absence of morphemes or syntactic features, and 2) the absence or lack of applicability of transformations which European languages have. We have also seen that some of the differences between Thai and more familiar languages which seem most profound are tracable to the most trivial of actual causes. If it turns out to be true, as some linguists are now arguing,²⁶ that the deepest level of language consists of semantic structures and it is the transformations alone which may properly be called syntax, grammarians like Lanyon-Orgill may be correct in a deeper sense than they ever dreamed of when they say that Thai does not possess any grammar in the European sense of the term.

4.0 Noun Phrase Structure. In addition to the comments above, it would be well to briefly outline the structure of Thai noun phrases, since these structures have the greatest bearing on our compounding study.

The basic structure of a Thai noun phrase is the head noun followed by modifiers. The modifiers are of three types: 1) relative clauses, 2) numbers, and 3) determiners. All types of modification which are not either numerical or determiners are derived from relative clauses. Relative clauses and determiners may, and numbers must occur with classifiers, or counters. Classifiers are nouns, generally of very general

meaning, which are used to count or precisely designate specific nouns. For example, the noun khon 'person' is the classifier for all nouns referring to people. Thus, in order to say "two men", one says:

phû·chǎ·y sǎ·ŋ khon "two men"
man two person

With determiners, the classifier occurs first:

phû·chǎ·y khon ní· "this man"
man person this

With adjectives (reduced from relative clauses), the classifier also occurs first:

phû·chǎ·y khon yà·y "the big man"
man person big

Classifiers may, but seldom do, occur with full relative clauses:

ʔǎ· kǎ· phû·chǎ·y khon thī· ye·ŋ phû·yin cà·k
well then man person who snatch girl from

khon takí· ŋay lǎ? khráp
person previous how (ptl.) sir

"Well then, he's that man who snatched the girl from the other fellow-- you know?"

When all three types of modification occur with the same noun, the relative clause is first, followed by the number and the determiner:

mǎ· (tua) lék sǎ·ŋ tua nán ²⁷
dog body small two body that

"Those two little dogs."

It is the relative clause modifiers which are involved in our analysis of noun compounds.

Footnotes

¹In addition a few phonological rules are tentatively proposed to account for alliterative, rhyming, and vowel and tone ablaut features in the so-called reduplicated forms.

²E.g. Noam Chomsky, Syntactic Structures (The Hague, The Netherlands: Mouton & Co., 1957) and Aspects of the Theory of Syntax (Cambridge, Mass.: The MIT Press, 1965).

³In fact the theoretical justification for a distinct semantic component as an interpreting device has been called into question. See chapter 2 for discussion. We will present a few phonological rules in the appendix but will withhold discussion until then.

⁴Gordon H. Allison, Modern Thai (Bangkok: Prachandra Printing Press, 1959), p. vii.

⁵Peter A. Lanyon-Orgill, An Introduction to the Thai (Siamese) Language for European Students (Victoria, B.C.: The Curlew Press, 1955), p. 35.

⁶Allison, loc. cit. On the same page, Allison states that "The writer has never yet met a Thai person who thinks that Thai grammar is easy." I am equally sure that no English speaking person would think English grammar was easy if its rules were "patterned to some extent" on rules of Far Eastern languages!

⁷Stuart Campbell and Chuan Shaweevongse, The Fundamentals of the Thai Language, second edition (Bangkok: Thai Australia Company, 1956), p. 142.

⁸Allison, loc. cit.

⁹Ibid., p. 7.

¹⁰Ibid., p. 10. Actually this statement is not quite true and its exceptions raise an interesting problem which is discussed in chapter 4.

¹¹Campbell and Shaweevongse, op. cit., p. 3.

¹²Mary R. Haas and Heng R. Subhanka, Spoken Thai (New York: Henry Holt & Co., 1945), p. 24, 25.

¹³Ibid.

¹⁴Edward M. Anthony et. al., Foundations of Thai, Book I, Part I (Pittsburgh: University of Pittsburgh, 1966), p. 1.

¹⁵The statement quoted above (pg. 4) from Allison concerning indefinite adjectives refers more to the use of indefinite determiners rather than their absence.

¹⁶Jerrold J. Katz and Paul M. Postal, An Integrated Theory of Linguistic Descriptions (Cambridge, Mass.: The MIT Press, 1964), p. 8.

¹⁷Richard B. Noss, Thai Reference Grammar (Washington, D.C.: Foreign Service Institute, 1964), p. 216, 217.

¹⁸Chomsky, Aspects, p. 171, 172 and note 27, p. 232.

¹⁹Ibid., p. 172 and note 35, p. 221.

²⁰But cf. the work of Fillmore (E.g. Charles J. Fillmore, "Concerning English Prepositions" in Francis P. Dineen, S.J. (ed.), Report of the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies, Monograph Series on Languages and Linguistics No. 19 (1966) (Washington, D.C.: Georgetown University Press)), in which case is viewed as more basic than tree structure.

²¹It is clear that these features are present; it is less clear that they have syntactic significance. However, it will be recalled that Thai has certain pronouns which are restricted to men or to women (See quotation from Haas and Subhanka, p.). I would maintain that the ungrammaticality of a sentence like *phī·cha·y phū·t wā· dīchan... "Big brother said, 'I...'" is based on restrictions between the feature [+Male] in cha·y and [-Male] in the pronoun dīchan 'I (woman speaking)'.²²

²²Chomsky, Aspects, p. 175.

²³Arguments for this analysis have been given by Lakoff (George Lakoff, "On the Nature of Syntactic Irregularity" (Cambridge, Mass.: The Computation Laboratory, Harvard University, Report No. NSF-16, 1965)) and Ross (John Robert Ross, "Adjectives as Noun Phrases" (Massachusetts Institute of Technology, mimeographed, 1966)).

²⁴Ross, op. cit., p. 3. I have simplified the structure somewhat.

²⁵But as we shall see in chapter 4, Thai has a structure similar to the one Ross gives for English from which sentences are developed in which the Thai copula pen does appear before adjectives, like in English. If Ross is right, then the difference between the two languages is even more trivial: this sort of development is obligatory in English, optional in Thai, but possible in both.

²⁶See the discussion of these matters in chapter 2.

²⁷This example is from Mary R. Haas, "The Use of Numerical Classifiers in Thai", Language XVIII (1942), 204. This article is a concise, but thorough, description of the use of classifiers.

Chapter 2

Input to Transformations

1.0 Introduction. In the history of transformational grammar, there have been a series of major revisions of the deepest part of grammar on which the transformations operate. The matter is far from settled, even at the present time. It is not our purpose, while dealing with an aspect of Thai grammar, to make any proposals with regard to this part of the theory or even to select among the candidates presently available, but to briefly examine the development of this part of the theory and to decide what structures must be available to the transformational component of a grammar of Thai which accounts for noun compounds. It will be of interest to notice the following factors at each level of development: 1) The nature of the rules, 2) the nature of the structures generated, 3) the context-sensitivity issue, 4) rule ordering, 5) universals, 6) peculiarities of each proposal. The varieties of basic systems which will be discussed are: Chomsky's in Syntactic Structures,¹ Katz and Postal's in An Integrated Theory of Linguistic Descriptions,² Chomsky's in Aspects of the Theory of Syntax,³ unpublished proposals of James McCawley, Stephen Anderson, and others in late 1966 and early 1967,⁴ and the most recent proposals of McCawley and others, dealing with grammar and logic.⁵

2.0 Syntactic Structures. To a degree, the organization of Syntactic Structures results in a certain emphasis on the phrase structure component which is the deepest level at this initial stage. Chomsky's approach is to present some theoretical grammatical models and to decide just what sort of model is necessary to deal with the grammar of natural languages in general and English in particular. A phrase structure grammar is thus shown to be superior to a finite state grammar, but not to be entirely adequate to deal with some aspects of real languages by itself. But when the notion of grammatical transformation is added, solutions to these problems are readily found. Apparently because of this approach, Chomsky does not begin discussing the relationship of the phrase structure and the transformations in the whole grammar. The net result is to make it appear that the transformations are an appendix, although an important, even crucial one, to the basic phrase structure grammar. Hence the phrase structure part of the grammar, which supplies structures to the transformations, enjoys an importance for syntax which it does not have in subsequent formulations.

The rules of the phrase structure section of the grammar are rewriting rules and are said to be presupposed by constituent analysis. The rules are of the form $X \rightarrow Y$ which is interpreted "as the instruction 'rewrite X as Y'".⁶ Sentences in languages are generated by a series of such rules called a derivation, such that each rule after the first rewrites a

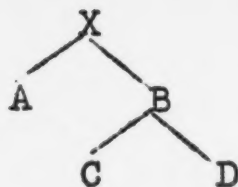
symbol which appears at the left of a rule which has previously applied. This process continues until no symbols are left but terminal symbols to which no further rules apply. Thus it is clear that the phrase structure is most basically a system which deals with strings of symbols.

A rewriting system is presupposed by constituent analysis and the other side of this coin is that every derivation can be represented "in an obvious way" by means of a tree diagram.⁷ To take a simple hypothetical example, the derivation:

$$X \longrightarrow A + B$$

$$B \longrightarrow C + D$$

may be represented by this tree:



Although "The diagram (15) conveys less information than the derivation (14)" yet "The diagram (15) retains just what is essential in (14) for the determination of the phrase structure (constituent analysis) of the derived sentence...".⁸

Since the diagram is not as explicit as the derivation, it is possible for a single diagram to correspond to more than one derivation. When this happens, the diagram represents a structurally ambiguous sentence, with as many readings as derivations. It is theoretically possible, then, for phrase structure grammars to generate ambiguous sentences. Chomsky leaves open the question whether or not the phrase structure component

of a grammar of any language actually does generate ambiguous sentences:

In fact, it is not clear that there are any cases of constructional homonymity purely within the level of phrase structure once a transformational grammar is developed.⁹

As long as this is in doubt, the derivation of strings of symbols, which does not admit of ambiguity, must be seen as more basic than the tree diagram, which does.

On the other hand, once transformations are admitted to the grammatical theory, it emerges that there are transformational rules which require more information for their application than that which is contained in strings of symbols, namely they require reference to the constituent structure of the strings to which they apply.¹⁰ So there are hints even in Syntactic Structures that the structures associated with derivations will ultimately prove more important than the strings of symbols which the derivations generate, and even the derivations themselves.

Restrictions on the structures (or strings) which the phrase structure rules generate must be drawn, so the issue of context sensitivity arises. There may well be cases in which $X \rightarrow Y$ is not generally applicable, but only when X occurs in the environment W_Z . In English, to take one of Chomsky's examples, the determiner symbol T may be rewritten a if the following noun is singular, but not if it is plural.¹¹ That is, a boy is well-formed, but *a boys is not. The rules in

the phrase structure section of a grammar must be context-sensitive in this sense. The above rule may be written:

$$X \longrightarrow Y \text{ in } W\_Z$$

which is to be interpreted as "the symbol X is to be rewritten as the symbol Y if and only if it is immediately preceded by W and immediately followed by Z".

Although the rules must be written on paper in some order, they can be regarded as ordered or unordered. If they are unordered, any rule can apply as soon as the symbol appearing in its right half has been generated. Suppose a phrase structure grammar contains the following three unordered rules:

$$1. A \longrightarrow B + C$$

$$2. B \longrightarrow D + E$$

$$3. C \longrightarrow F + G$$

Then, as soon as 1. has applied both B and C are available and either 2. or 3. can apply in either order. On the other hand, if the rules are ordered, 2. must apply before 3. regardless of the fact that D has already been generated after applications of 1. Chomsky asserts that ordered rules are desirable:

The formal properties of the system of phrase structure make an interesting study, and it is easy to show that further elaboration of the form of grammar is both necessary and possible. Thus it can easily be seen that it would be quite advantageous to order the rules of the set F so that certain of the rules can apply only after others have applied.¹²

The idea that linguistic theory should be concerned with the universal characteristics of language is present in

Syntactic Structures

More generally, linguists must be concerned with the problem of determining the fundamental underlying properties of successful grammars. The ultimate outcome of these investigations should be a theory of linguistic structure in which the descriptive devices utilized in particular grammars are presented and studied abstractly, with no specific reference to particular language.¹³

But there are no specific proposals of particular characteristics of these universal descriptive devices except for the general form of grammar itself.

From these considerations we are led to a picture of grammars as possessing a natural tripartite arrangement. Corresponding to the level of phrase structure, and corresponding to lower levels it has a sequence of morphophonemic rules of the same basic form. Linking these two sequences, it has a sequence of transformational rules.¹⁴

Probably the best-known particular characteristic of Syntactic Structures is the notion of kernel sentence. This idea, which captured the imagination of people interested in applied linguistics, did not survive later revisions of the theory. A kernel sentence is one of:

the set of sentences that are produced when we apply obligatory transformations to the terminal strings of the $[Z, F]$ grammar. The transformational part of the grammar will be set up in such a way that transformations can apply to kernel sentences (more correctly, to the forms that underlie kernel sentences - i.e., to terminal strings of the $[Z, F]$ part of the grammar) or to prior transforms. Thus every sentence of the language will either belong to the kernel or will be derived from the strings underlying one or more kernel sentences by a sequence of one or more transformations.¹⁵

The seeds of the demise of kernel sentences are contained in the parenthetical sentence in the above quotation; further study of these "underlying forms" showed that they had little to do

with kernel-type sentences.

3.0 Katz and Postal. Another step in the development of generative-transformational grammar is represented by Jerrold Katz and Paul Postal's work in An Integrated Theory of Linguistic Descriptions.¹⁶ Katz and Postal were interested in providing a theory which would deal with the relationship between grammar and meaning and this led to a number of modifications of the notion of the phrase structure part of a transformational grammar. The rules are essentially the same in nature as they are in Syntactic Structures, i.e. they operate on strings of symbols by a rewriting process.¹⁷ Presumably the rules are still context-sensitive and ordered, although Katz and Postal do not discuss these matters, but acknowledge that Chomsky's formulation is correct.¹⁸

As in Syntactic Structures, tree structures are associated with the strings of symbols produced by the rewriting rules. But these associated trees, or labeled bracketings, are seen as providing formalizations which are not readily available in the derivations themselves:

Such rules permit the construction of derivations: finite sequences of strings of symbols, beginning with the initial sequence of the grammar #Sentence#... There is an algorithm or mechanical procedure for associating a labeled bracketing or P-marker with each such derivation... Such labeled bracketings formally render the notions of grammatical category, part of speech, or immediate constituent structure.¹⁹

Furthermore, it is on these structures, and not on kernel sentences, or even the symbols underlying them, that the

transformations operate.²⁰ The transformations take existing P-markers and modify them by adding, moving, or deleting branches. Thus it can be seen that the labeled bracketing is more important in the theory than in the previous version. In addition, the phrase structure component is so constructed that all information necessary to determine the meaning of a sentence is present in the underlying P-marker and that the application of transformations, at least singular transformations, does not affect meaning.²¹ Earlier, it was assumed that the difference in meaning between questions and statements, for example, was accounted for in connection with the operation of the transformation which converted statements into questions. Thus questions and statements had the same underlying string and the same underlying P-marker. But if all the information necessary to account for the meaning of a sentence is to be present in the underlying phrase marker and questions and statements differ in meaning, it follows that this state of affairs cannot continue. Katz and Postal solve this problem by positing a morpheme Q (Question) which appears in the underlying P-markers of questions but not in the P-markers of the corresponding statements. The semantic rules are constructed to read question meaning from this morpheme. Furthermore, the question transformation is now obligatory; it must apply to all P-markers containing Q. Q itself does not appear in the structure of the sentence as it is spoken (the final derived P-marker), but is converted into question intonation.²² Similar

solutions for imperative, negative, and passive sentences are posited.

In An Integrated Theory of Linguistic Descriptions, we find a few definite claims as to what some universals are. One of these has to do with embedding. In the underlying P-marker of every sentence in which another sentence is embedded, there is a position marked by a matrix dummy symbol. At this point, another P-marker will be embedded by the embedding transformation. Among the symbols which dominate these matrix dummies are two which are universal.

We specify further that all syntactic components contain among the nonterminal symbols of their phrase structure subpart a specified set of constituents including at least two, called Relative (Rel) and Complement (Comp).²³

The presence of universal symbols implies the presence of universal phrase structure rules:

Besides assuming a certain universal grammatical vocabulary, we assume also a certain set of universal phrase structure rules. We claim that the grammars of all languages introduce elements like Rel and Comp as subparts of the major constituents like Noun Phrase, Verb Phrase, etc. In other words, elements like Noun Phrase and Verb Phrase will dominate, among other things, sequences of universal elements like Rel and Comp...²⁴

If one were to select a distinctive feature of the phrase structure component as explicated in An Integrated Theory of Linguistic Descriptions, it would be the idea that underlying P-markers contain all the information necessary to provide the meaning of the sentence.

4.0 Aspects of the Theory of Syntax. Further

modifications of the theory appeared in the work of Chomsky in Aspects of the Theory of Syntax. The rules proposed in this work include some of completely different nature. The basic type are symbol rewriting rules of the previous type. Although the rules rewrite symbols, they are said to be "a natural mechanism for generating Phrase-markers."²⁵ The generation of Phrase-markers appears more important than the generation of strings of symbols, although the Phrase-marker is still assigned on the basis of the derivation, which is written in terms of symbols.²⁶ But except for this rather subtle shift in emphasis, the nature of the rules and the method by which structures are assigned to strings remains essentially the same. However, a new type of rule is added to the familiar type. First, there are rules which analyze lexical categories, like N(oun) and V(erb) into complex symbols. These complex symbols are sets of specified syntactic features, analogous to phonological features.

For example, we might have the following grammatical rules:

- (20) (i) $N \Rightarrow [+N, \overset{+}{-}Common]$
 (ii) $[+Common] \Rightarrow [\overset{+}{-}Count]$
 (iii) $[+Count] \Rightarrow [\overset{+}{-}Animate]$
 (iv) $[-Common] \Rightarrow [\overset{+}{-}Animate]$
 (v) $[+Animate] \Rightarrow [\overset{+}{-}Human]$
 (vi) $[-Count] \Rightarrow [\overset{+}{-}Abstract]$ ²⁷

Various choices represented by the $\overset{+}{-}$ symbols in the above rules allow the generation of various complex symbols for the

lexical symbol N. Rules such as those of (20) in the citation above are called subcategorization rules, since they express subcategories of the lexical category Noun. Subcategorization rules need not be context-free, as those in the example, but may also be context-sensitive. For example, a transitive verb would be specified by the feature [+__NP], meaning it can appear before a Noun Phrase. This feature is not derived from a rule like those above but gets it by convention from the fact that Noun Phrases follow Verb complex symbols when the verb must be transitive. If a context-sensitive subcategorization rule derives its syntactic features from categories like NP, it is called a strict subcategorization rule. Another type of context-sensitive subcategorization rule specifies complex symbols in terms, not of category symbols, but of the syntactic features of the environment. These rules are called selectional rules. Chomsky's examples of strict subcategorization rules and selectional rules, respectively, are:

$$(40) V \Rightarrow CS/ \left\{ \begin{array}{l} NP \\ \# \\ Adjective \\ Predicate-Nominal \\ like \sim Predicate-Nominal \\ Prepositional-Phrase \\ that \sim S' \\ NP (of \sim Det \sim N) S' \\ etc. \end{array} \right\}$$

and:

$$(42) \left\{ \begin{array}{l} (1) \\ (11) \\ (111) \\ (1v) \end{array} \right\} [+V] \Rightarrow CS/ \left\{ \begin{array}{l} [+Abstract] Aux \_ \\ [-Abstract] Aux \_ \\ \_ Det [+Animate] \\ \_ Det [-Animate] \end{array} \right\} 28$$

If a verb complex symbol appears in the first categorial environment of rule (40), it receives the feature [+__NP] (i.e. is transitive). If it appears in the second environment, it receives the feature [+__#] (i.e. is intransitive), and so on. The rules in (42) are designed to select verbs in terms of the features of other lexical categories in their environments, e.g. (i) selects verbs which take Abstract subjects, (ii) selects verbs which take Concrete subjects, (iii) selects those which require Animate objects, and so on.

It can be seen that the subcategorization rules create complex symbols which are made up of various syntactic features necessary to assure that the sentences of the language not only have the right constituent structure (which is provided by the rewrite rules of the older type) but consist of strings of words which are grammatically compatible with each other. The actual lexical items of the language are then matched with the appropriate complex symbols by a general lexical rule. The lexicon is a set of entries which each contain sets of syntactic, semantic, and phonological features idiosyncratic to the individual entry. The general lexical rule allows a lexical item to replace a complex symbol if its inherent features are not distinct from the features of the complex symbol. Thus, since the verb eat in English has the inherent syntactic features [+__NP] and [+Animate Aux__], it may replace a verb complex symbol only if it too has these features. In this way, appropriate lexical items appear in the string which

will ultimately be a grammatical sentence in the language.

As an alternative to the above process, Chomsky proposes that complex symbols be eliminated and that the lexical items be introduced into structures by a convention which says, in effect, that a lexical entry may be entered into a structure unless it has in its syntactic feature set features which restrict it from entering such structures. If a lexical item has a feature [+X__Y], it may replace a symbol of the preterminal structure if that symbol actually does appear in the environment X__Y and not otherwise. The two methods of introducing lexical items into sentences are not exactly equivalent, but Chomsky does not decide between them.²⁹

While the structures generated by this version of the base component are essentially the same as those of the earlier versions, there is one important difference.³⁰ Generalized transformations which take two Phrase-markers and combine them are no longer part of the theory. Instead, the symbol S(entence) can now appear to the right of rules in the base component. In this way, Sentence nodes can be embedded under nodes which are themselves embedded under other Sentence nodes. As a result, embeddings are taken care of in the base component instead of in the transformations.³¹

The facts previously taken care of by requiring that the phrase structure rules be context-sensitive can now be handled by the subcategorization rules, or by the "filter effect" of the transformations. Since S can now appear to the right of

base component rules and can be subsequently developed in any way any other S in the language can be developed, there is no way to assure that it will be compatible with the sentence in which it is embedded. However, the transformations which carry out the various deformations of embedded sentences so as to produce grammatical sequences will only operate in those cases in which the embedded and embedding sentences are compatible. So we assume the convention that any structure generated by the base component which has failed to undergo the appropriate transformations is not a real sentence in the language. As a result, the transformations serve to assure compatibility in the case of embedding. As we have seen, the subcategorization rules are designed to assure compatibility within single sentences. Between the "filter effect" and the subcategorization rules, then, the need for context-sensitive branching rules has been eliminated. As Chomsky puts it:

We suggested: (a) that the distributional restrictions of lexical items be determined by contextual features listed in lexical entries, and (b) that these contextual features be regarded as defining certain substitution transformations. Thus strict subcategorical and selectional restrictions of lexical items are defined by transformational rules associated with these items. We have now observed that the transformational rules must also carry the burden of determining the distributional restrictions on base Phrase-markers. Thus the categorial rules that generate infinite sets of generalized Phrase-markers can apparently be context-free, with all distributional restrictions, whether of base Phrase-markers or lexical entries, being determined by the (singular) transformations.³⁴

In Aspects, Chomsky is quite interested in the subject of universals, i.e. those aspects of grammar which are common

to all languages. Universals may be either formal (pertain to the general form which a grammar of any natural language must have) or substantive (pertain to specific details which are characteristic of all languages). It has been the clear goal of transformational linguistics from the beginning to explore formal universals, but until recently, little attention has been devoted to substantive universals. As we have seen, Katz and Postal suggest that some details in the base component may be universal. But Chomsky in Aspects begins to hint that much more than a few details may be involved:

To say that formal properties of the base will provide the framework for the characterization of universal categories is to assume that much of the structure of the base is common to all languages.³³

But Chomsky stops short of saying just what parts of the structure of the base is universal or to outline anything which might claim to be that universal base.

There is little doubt that the notions involved in the explication of the subcategorization rules are the most prominent feature of the base component at this stage. It is also interesting to note that the notion of kernel sentence, which was distinctive in Syntactic Structures, all but disappears in Aspects:

Among the sentences with a single base phrase-marker as basis, we can delimit a proper subset called "kernel sentences." These are sentences of a particularly simple sort that involve a minimum of transformational apparatus in their generation. The notion "kernel sentence" has, I think, an important intuitive significance, but since kernel sentences play no distinctive role in generation or interpretation of sentences, I shall say nothing more about them here.³⁴

5.0 The Base Component of a Transformational Grammar.

The version of the base component discussed by McCawley in 1966 and Anderson involves several other departures from what has gone before. The rules in the McCawley paper are not rewriting rules at all. What corresponds to rules of the rewriting type in the earlier versions is now a set of conditions of well-formedness on underlying phrase-markers in the base component. This differs from the rules in Aspects in a number of important ways. In the first place there is no provision for rewriting strings of symbols; strings of symbols play no significant role whatsoever. For this reason the earlier notion of an algorithm or mechanism for associating structures with derivations disappears. The set of conditions are to define, not produce, the needed structures. Finally, the conditions are unordered, in contrast to the ordered rewrite rules.³⁵

In Anderson's own work, he proposes that the set of conditions stated in McCawley's notation be replaced with a set of other conditions having to do with the domination of one node by another. These conditions are even less like rules than are McCawley's, although they define the same (universal) set of trees, with the exception that they say nothing about left-to-right ordering of the nodes. This ordering is to be provided by a set of rules, differing for each language, which specify the order of nodes for that language.³⁶ However, although the relationship between the rules or conditions and

the structures is markedly different in this type of formulation, the structures are the same in nature, only simpler in detail and more explicitly defined.

Little is said here about the problems of context-sensitivity, but, for Anderson at least, they apparently would be handled in a way similar to that proposed by Chomsky in Aspects.

In this conception of grammar, the base would consist of a lexicon with appropriate insertion rules for the attachment of lexical items to the terminal nodes of underlying constituent structure trees in accordance with the selectional and subcategorizing features peculiar to the individual lexical items...³⁷

Strong claims concerning the universality of the base component are made in these papers. It is suggested that it is possible to specify just the structures which appear in the deep structures of any language:

Recent research in generative grammar has led to a detailed examination of the possibility that some significantly large portion of the base component is not a part of the grammars of particular languages at all, but is rather provided by universal theory. The strongest hypothesis about the contribution of universal grammar to the base so far taken seriously is the assertion that the categorial component of the grammar consists of exactly the following unordered conditions...³⁸

Following this statement are seven sets of specifications on structures. The details need not concern us here. The point of interest is that there is a growing search for a statement of substance concerning a putative universal base component. Anderson goes on to reinterpret the conditions, and to further restrict the amount attributable to the universal base, but retains the idea of such a base.

This notion of a universal base and the elimination of the notion of rewriting rules are radical innovations at this stage of development. In addition, Anderson proposes a set of principles for the labeling of nodes up in a tree, given a tree structure and lexical category symbols. These principles apply not only to labeling in the base component, but also to the relabeling of trees as they are deformed by transformations. We shall examine these principles in detail in connection with our discussion of the transformations necessary to generate Thai noun compounds.

So rapid has been the recent development of conceptions at this level of grammar, that Anderson now disclaims much of the detail in his discussion.³⁹ McCawley has made a new proposal, which we shall now examine.

6.0 "Logical Grammar". In the previously discussed version of the part of grammar providing the input to transformations, rewriting rules were abolished. But a shadow of the concept of rule remained in the form of specific conditions on structures in the universal base further modified by language-particular restrictions on basic structures. In the latest proposal of McCawley the notion of rule seems to disappear altogether. In fact, the base component itself is taken to be a superfluous level which should be eliminated. Instead, we need deal only with a (presumably largely universal) set of possible messages and a code to relate these to the surface forms of sentences.⁴⁰ It is the domain of

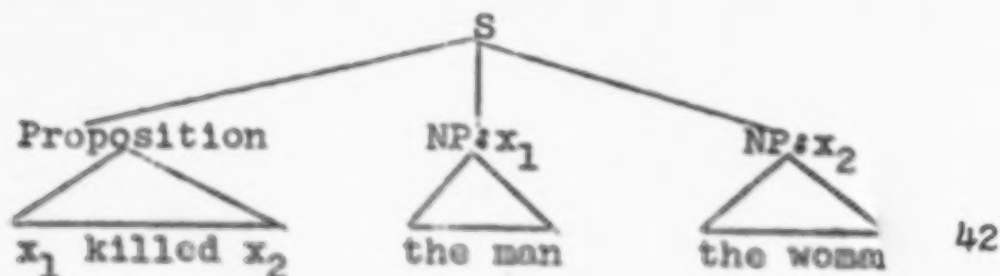
semantics to deal with what is or is not a possible message and of the transformations to relate these possible messages to sentences in languages. Between, there is no need for another level corresponding to the base component.

Nevertheless, there is a certain type of structure appropriate to express the range of possible messages. A modified form of symbolic logic is proposed as the framework of this expression:

I will in fact argue that symbolic logic, subject to certain modifications, provides an appropriate system for semantic representation within the framework of transformational grammar.⁴¹

In particular:

I conclude that it is necessary for semantic representation to separate an utterance into a 'proposition' and a set of noun phrases, which provide the material used in identifying the indices of the 'proposition', e.g.



The triangles here (and for that matter throughout this present paper) are to be interpreted as indicating that the details of structure at this point are not given. Clearly, structures of the above type are not very much like those generated by earlier phrase structure and base components.

In discussing context-sensitivity, McCawley argues that the cases handled by Chomsky's strict subcategorization rules have to do with logical well-formedness and not grammar.

Some sentences which would be in violation of selectional restrictions should not be ruled out at all, since there are situations in which they are, indeed, possible messages. While a sentence like "My toothbrush is trying to kill me." seems deviant, it is perfectly reasonable as a report of a dream ("I dreamed that my toothbrush was trying to kill me."), or as a statement of the beliefs of someone else ("John thinks his toothbrush is trying to kill him.")⁴³ In other cases, selectional restrictions exclude sentences that should be excluded. Activity verbs, for example, surely must be prevented from taking the be progressive, since there is no: *"I dreamed (or Arthur believes) that John was knowing the answer."⁴⁴

More crucially, this new concept of the nature of language deals with a sort of "context-sensitivity" which no previous model has been able to handle at all. There are numerous cases, especially in embeddings, when it is necessary for there to be a system of cross-referencing between elements in different parts of sentences. McCawley discusses in detail a number of such cases which cannot be dealt with successfully by means of a grammar which depends on a base component.⁴⁵

Although he does not deal with universals in detail, it seems clear that many of the restrictions on 'possible messages' will be universal, while others may be language-particular. The interest in universals is certainly no less strong than in his 1966 proposals.

The notion of the role of the principles of symbolic

logic in the deepest aspect of language and the elimination of the base component are the most distinctive features of this latest proposal.

While it is not the purpose of this paper to decide on the theoretical issues raised in this discussion, it is necessary to specify the structures which are available to the transformations in Chapter 3. The more recent work has provided cogent reasons for regarding rewriting systems as too elaborate for this purpose. The most recent work of McCawley⁴⁶ is still in its initial stages of development, and must in any case assume an early set of transformations to convert the modified symbolic logic structures into trees of the more familiar type, unless the transformational component is to be radically revised. So we will state the conditions which specify the structures we will need after the manner of Anderson's paper, assuming that these will also appear at some stage in a grammar based on more recent notions. We use the symbolization of McCawley's earlier paper.⁴⁷ The conditions are:

<S; S*>

This means that a sentence may consist of any number of conjoined sentences.

<S; NP VP>

A sentence may also consist of a noun phrase and a verb phrase.

<NP; N(S) (No) (Det)>

A noun phrase consists of a noun and optionally a sentence.

a number, and a determiner, in that order.

<NP; S>

A noun phrase may also be a sentence.

<VP; V (NP) (NP)>

A verb phrase consists of a verb and optionally, one or two noun phrases.

These conditions differ from the proposed universal set in Anderson's paper in several ways. First, there are no conditions providing for noun phrase and verb phrase conjunction apart from sentence conjunction in our set of conditions. All cases of conjunction of noun phrases which we will deal with can be assumed to result from conjunction of full sentences. There are three specifications above which differ from those in Anderson's list. Our condition <NP; N (S) (No) (Det)> corresponds to Anderson's <NP; N (S)>. It is clear that noun phrases containing adjectives are developments from noun phrases with embedded sentences, but such a solution will not account for noun phrases with numbers and determiners.⁴⁸ On the other hand, our <NP; S> is simpler than his <NP; NP S>. Anderson does not say, but I surmise that the presence of NP is for the purpose of dominating certain instances of it which have been posited in the deep structure of English and other languages to account for sentences like: "It surprises me that the doctor came at all."⁴⁹ There are no such sentences in Thai and therefore no need for deep structure t. Finally, our <VP; V (NP) (NP)> lacks the (S) of Anderson's <VP; (NP) (S)> and includes another optional NP. There are no cases of

verb phrase complementation in this presentation of Thai grammar.⁵⁰ The second NP seems to be needed for structures which will prove important to us, namely the verb phrase meaning "to use something for something". The verb phrase specification above gives this structure two noun phrases and the preposition meaning "for" is inserted by the rule described below.

These conditions, furthermore, are only adequate if the following transformations, which will not be discussed in detail, are assumed: (1) A rule which places prepositions in the phrase markers of verb phrases, based on syntactic features possessed by certain verbs by a "segmentalization" process. In this way, the preposition kàp 'with' will be segmentalized from features of the main verb to generate kíaw kàp 'to deal with'. (2) A rule which generates classifiers, whose behaviour is discussed in Chapter 1, from features of nouns. (3) A transformation, probably universal, which accounts for the insertion of the conjunction lě? 'and' in the appropriate structures. (4) A (perhaps universal) rule of conjunction reduction which reduces repeated material in conjoined structures. More about this will be said in Chapter 7. (5) Finally, and this has nothing to do with transformations, we assume that one verb at least is marked as the copula and that its presence leads to copula structure readings.⁵¹

If these conditions are met, we have the structures necessary to provide the input to the transformations in Chapter 3.

Footnotes

¹Chomsky, Syntactic Structures.

²Jerrold J. Katz and Paul M. Postal, An Integrated Theory of Linguistic Descriptions (Cambridge, Mass.: The M.I.T. Press, 1963).

³Chomsky, Aspects.

⁴James D. McCawley, "Concerning the Base Component of a Transformational Grammar" (unpublished, 1966). Stephen R. Anderson, "Concerning the Notion 'Base Component of a Transformational Grammar'" (unpublished, 1967).

⁵James D. McCawley, "Where Do Noun Phrases Come From?" (to appear in Roderick Jacobs and Peter S. Rosenbaum (eds.), Readings in English Transformational Grammar). James D. McCawley, "The Respective Downfalls of Deep Structure and Autonomous Syntax" (Paper read at the 42nd Annual Meeting of the Linguistic Society of America, Chicago, Illinois, December 30, 1967). D. Terrence Langendoen, "On Selection, Projection, Meaning, and Semantic Content" (to appear in Working Papers in Linguistics, Ohio State University Research Foundation).

⁶Chomsky, Syntactic Structures, p. 26.

⁷Ibid., p. 27.

⁸Ibid., p. 27, 28.

⁹Ibid., p. 87, footnote 2.

¹⁰Ibid., p. 40, 43.

¹¹Ibid., p. 28.

¹²Ibid., p. 33.

¹³Ibid., p. 11.

¹⁴Ibid., p. 45, 46.

¹⁵Ibid.

¹⁶Katz and Postal, op. cit.

¹⁷Ibid., p. 6.

¹⁸Ibid.

¹⁹Ibid.

²⁰Ibid.

²¹Ibid., p. 32, 33. A suggested method for handling linguistic facts previously handled by generalized transformations so that there are no transformations at all which affect meaning is given on p. 67, 68.

²²Ibid., p. 110, 111.

²³Ibid., p. 48.

²⁴Ibid.

²⁵Chomsky, Aspects, p. 66.

²⁶Ibid.

²⁷Ibid., p. 82.

²⁸Ibid., p. 94.

²⁹Ibid., p. 121-123. We shall handle context-sensitivity more or less in the method outlined here. We do not choose between the two alternatives in any meaningful sense, but, for the sake of ease of exposition, we shall write our descriptions as if the second alternative had been chosen.

³⁰There is no longer a phrase structure component, the name is now "base component". A phrase structure grammar (p. 88) consists of an unordered set of rewriting rules. Ordering the rules, hinted at in Syntactic Structures, was a serious modification from phrase structure grammar and the introduction of complex symbols is a radical departure.

³¹Chomsky, Aspects, p. 134. Examples of such rules occur on p. 102, 107, passim.

³²Ibid., p. 139.

³³Ibid., p. 117.

³⁴Ibid., p. 18.

³⁵Anderson, et al., p. 1.

³⁶Ibid., p. 4, 5.

³⁷Ibid., p. 4.

³⁸Ibid., p. 1.

³⁹In a cover letter sent with a copy of "Concerning the Notion 'Base Component of a Transformational Grammar'", Anderson says, "Like most specific proposals in linguistics, its content has evaporated fairly quickly, and I would no longer defend much of the detailed apparatus proposed in it."

⁴⁰McCawley, "Noun Phrases". Cf. Langendoen, op. cit., section 2.

⁴¹McCawley, "Noun Phrases".

⁴²Ibid.

⁴³Ibid.

⁴⁴Ibid.

⁴⁵Ibid. and McCawley, "The Respective Downfalls".

⁴⁶And others. See references in McCawley, "Noun Phrases".

⁴⁷McCawley, "Concerning the Base Component".

⁴⁸It may be that determiners are to be provided for by language-particular, not universal conditions, since not all languages have determiners. If so, our formulation does not contradict the one in Anderson's paper. But it seems entirely likely that numbers will have to be handled by universal constraints. Numbers present a number of interesting grammatical problems, which are too often ignored.

⁴⁹See for example Peter S. Rosenbaum, The Grammar of English Predicate Complement Constructions, (Cambridge, Mass.: The M.I.T. Press, 1967), pp. 11-13, passim. Interestingly, McCawley claims that his new formulation would eliminate the necessity for this it in the deepest level of language ("Noun Phrases").

⁵⁰Cf. Rosenbaum, op. cit., p. ix, where Rosenbaum reports that further work has cast doubt on the existence of verb phrase complementation in English.

⁵¹The Thai word pen 'to be, exist' is surely a copula. khǎi 'to be as follows' used to introduce lists, and chǎi 'to be named' may also be marked for this feature.

Chapter 3

Transformations

1.0 Introduction. Converting the structures specified in chapter 2 into the grammatical structures observable in the language as it is actually used is the function of the transformational component. Transformations perform the operations of addition, deletion, and rearrangement in carrying out its function.

Because the structures specified in chapter 2 allow sentences to be embedded in noun phrases, it is possible for a sentence to have any number of other sentences embedded in its constituents. In dealing with this feature of languages, the transformations apply cyclicly; that is, the same rules apply repeatedly over different domains. Specifically, the rules all apply to the innermost embedded sentences first, then to the next outer sentence, and so on until the structure is exhausted.¹

The transformations must be ordered with respect to each other. Sometimes in linguistic descriptions, it can be shown that a certain set of transformations are strictly ordered. This means that the rules must apply in exactly the order presented or certain unnecessary complications result. In other descriptions, rules can be shown to be partially ordered. That is, certain rules must occur in a certain order with

regard to some other transformations, but order does not matter in the case of others. Rules 4.1-4.15, with the exception of 4.8, are strictly ordered. The rest are partially ordered. As we describe the rules, we will specify which ones are effectively ordered and explain the undesirable results which would arise if another order were given.

We assume that the "filter" effect described in chapter 2 is a feature of transformations. Any structure which does not underlie any sentence in the language will not meet the structural conditions for any of the transformations in that language. If a sentence structure has an embedding to which no transformation applies, it must be rejected by the grammar. Chomsky has suggested that a sentence could fail to be acceptable by having in its structure embedded sentences surrounded by sentence boundaries (#S#).² As part of their operation, all transformations dealing with embeddings would delete these boundaries. Then, structures which are candidates for acceptable surface structures would have embedded sentences surrounded by sentence boundaries just in case no embedding transformations had applied to them. We assume that some such process is at work in our Thai grammar.

From Katz and Postal³ we borrow the principle that Pro-forms are freely deletable. In the process of compounding, as we shall see, certain constituents of a sentence are preserved in the compound and others are deleted. We have specified that the deleted noun forms are always Pro-nouns.

While they don't say so explicitly, the Pro-forms used by Katz and Postal are almost all nouns of very general meaning preceded by the indefinite determiner 'some'. Pro-forms, then are words like 'someone', 'something', 'someplace', etc. In Thai, the Pro-nouns are a subset of the classifiers, which are themselves a subset of nouns which have general meanings. An example of a Thai Pro-noun is khon 'person, people'. Nothing corresponding to the English determiner 'some' is necessary in Thai. Unlike Katz and Postal, we treat the notion Pro as a feature ([+Pro]) associated with items like khon rather than as a category symbol dominating them.

Because transformations add, delete, and rearrange, it sometimes happens that node labels become inappropriate during the operation of the transformations. Anderson proposes that certain principles of node labeling apply throughout the operation of the transformations.⁴ If at any point a node should happen to bear a label which violates one of these principles, it should be changed so as to conform to them. The principles he proposes are the following: (1) if a node dominates a node labeled N(oun), it is labeled N(oun) P(hrase), (2) if a node dominates a node labeled V(erb), it is labeled V(erb) P(hrase), (3) if a node dominates two nodes, one of which is labeled NP and the other is labeled VP, it is labeled S(entence), (4) if a node dominates only nodes of the same type (possibly except for a conjunction), it is a node of that

type, and (5) if a node exhaustively dominates another node of the same category, it is equal to that node (i.e. the upper of the two is deleted). If these principles apply throughout, superfluous and wrongly labeled nodes will be eliminated.

2.0 The Justification for Transformations. In dealing with transformations, one might ask what the basic function and purpose of transformations is. In other words, why should languages have transformations at all? If all the information necessary to understand a sentence is present in the deepest structures, why are not these simply converted directly into phonological material? Langendoen suggests that one basic answer to this question is that transformations in some cases render the deep structure more accessible to the language user than the direct conversion into phonological material would.⁵ It is a well-known fact that multiple layers of self-embedding render grammatical sentences extremely difficult to understand. Thus, to use one of Langendoen's examples, the sentence "The rumor that that report which the advisory committee submitted was suppressed is true is preposterous." is very hard to comprehend, although it is close to being a direct conversion of the deep structure. But if a transformation which moves embedded sentences to the end of the sentence in which they are embedded is repeatedly applied, the situation improves. Repeated application of this "extraposition" transformation gives "The rumor is preposterous that it is true that

the report was suppressed which the advisory committee submitted". This is much more easily understood, i.e., its deep structure is more accessible to the hearer.

Recalling the example of the completive aspect sentences in Thai and English which we discussed in chapter 1, we noted that the Thai example was closer to its deep form than its English counterpart. English transformations had moved the tense and aspect verbs, while no such changes had to take place in Thai. But recall that the English sentence involved a sentence which was a noun phrase which was a sentence whose verb was an aspect marker and this whole sentence was, in turn, a noun phrase whose predicate was a tense marker. The Thai sentence involved only one degree of this type of embedding, since it has no tense verbs. This notion of transformations as rendering deep structures accessible by reducing multiple embeddings may be a partial explanation of the difference between the two languages.⁶

But it seems to me that transformations perform another, partially conflicting function. One of the effects of transformations is to reduce the rather high degree of redundancy in deeper structures. To cite just two kinds of examples, in the cases of conjunction reduction and redundant noun deletion, transformations seem to have this function as a main purpose. Whatever the details, all the various schemes of conjunction reduction which have been proposed all involve elimination of repeated material in one of two conjoined structures. The

reduced form of the conjoined sentence "John went to the store and Harry went to the store" is "John and Harry went to the store" in which one occurrence of "went to the store" has been deleted. Similarly, the rule we call "redundant noun deletion" and which has been called "equi NP deletion" or "identity erasure" in work on English has as its function the deletion of a noun or noun phrase in an embedded sentence if a noun or noun phrase which has the same referent is present in the embedding sentence.

Perhaps the purpose of transformations in the most general sense is to add efficiency to the use of language as a vehicle of communication. Rendering deep structures accessible and reducing redundancy may well be two special cases of this general function.

3.0 Conventions. Certain conventions in symbolization will be used throughout the chapter. The following list explains them:

$\langle A \rangle_B$ A is a B; i.e. A is a constituent dominated by a node labeled B.

$\begin{bmatrix} +A \\ +B \end{bmatrix}$ used to designate a form with the grammatical features +A and +B.

$\begin{Bmatrix} A \\ B \end{Bmatrix}$ means a choice; either A or B, but not both.

(A) A is optional.

$A_1 \dots A_1$ Both A's share an identity of reference. For example, if the first A is the noun phrase "the

book", the second A not only is the noun phrase "the book", but refers to the same book.

R,U,W,X,Y,Z are used to designate any constituent whatsoever which is generable in the position in which they appear, including null.

Model rule form: $\langle A \ B \ C \ X \rangle_{NP}$

obl $\begin{matrix} 1 & 2 & 3 & 4 \\ 3 & 2 & \emptyset & 4 \end{matrix}$

A noun phrase consisting of the constituents A followed by B followed by C followed by any possible, or no constituent is altered by the deletion of A and the placement of C in the A position. The rule is obligatory. No symbol over the arrow indicates an optional rule.

4.0 The Rules. The rest of chapter 3 is a description of the transformations needed to generate the structures directly underlying the compounds themselves -- and a number of other Thai constructions as well. No attempt at exhaustiveness or comprehensiveness has been made. The few rules in this chapter are only those necessary to provide the basis for compounding. The fact that they also account for other constructions stems from the tendency of languages to have maximally general rules.

4.1 The Ambivalent Verb Transformation.

T Am Vb $\begin{bmatrix} +N \\ +Pro \end{bmatrix}$ $\langle \begin{bmatrix} +V \\ +Amb \end{bmatrix}$ $\langle X \begin{bmatrix} +N \\ \sim Hom \end{bmatrix} Y \rangle_{NP} \rangle_{VP} Z$

$\rightarrow \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 3 & \emptyset & 5 & 6 \end{matrix}$

This rule operates on sentences of which the main verb is an ambivalent verb. It deletes the subject if it is a Pro-noun, replacing it with the object of the ambivalent verb, provided that the object is non-human. The application of this rule leads to constructions which often must be translated passively in English, e.g.

rót khă·y nay krunthê·p
car sell in Bangkok

"Cars are sold in Bangkok."

derived from:

khon khă·y rót nay krunthê·p
person sell car in Bangkok

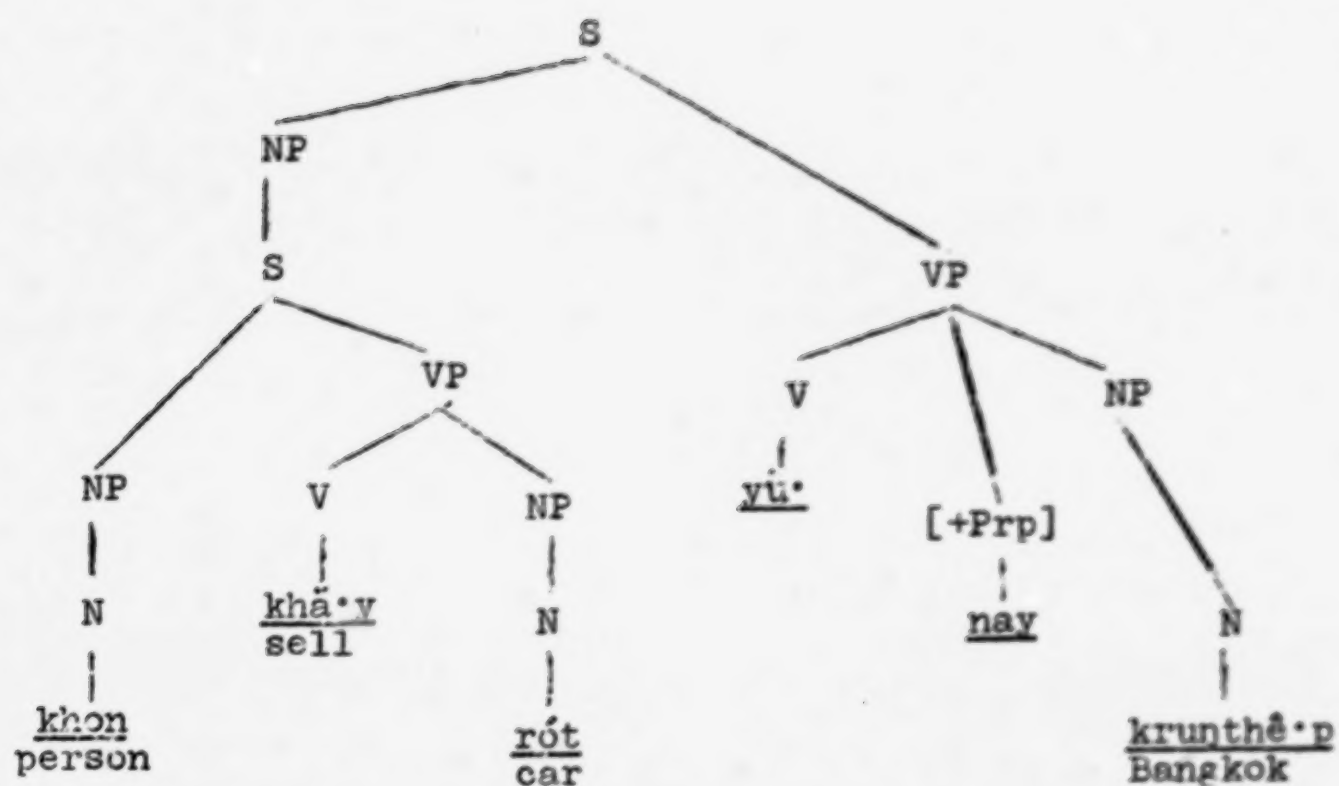
"They sell cars in Bangkok."

The operation of T Am Vb on these sentences is illustrated in Figure 6.

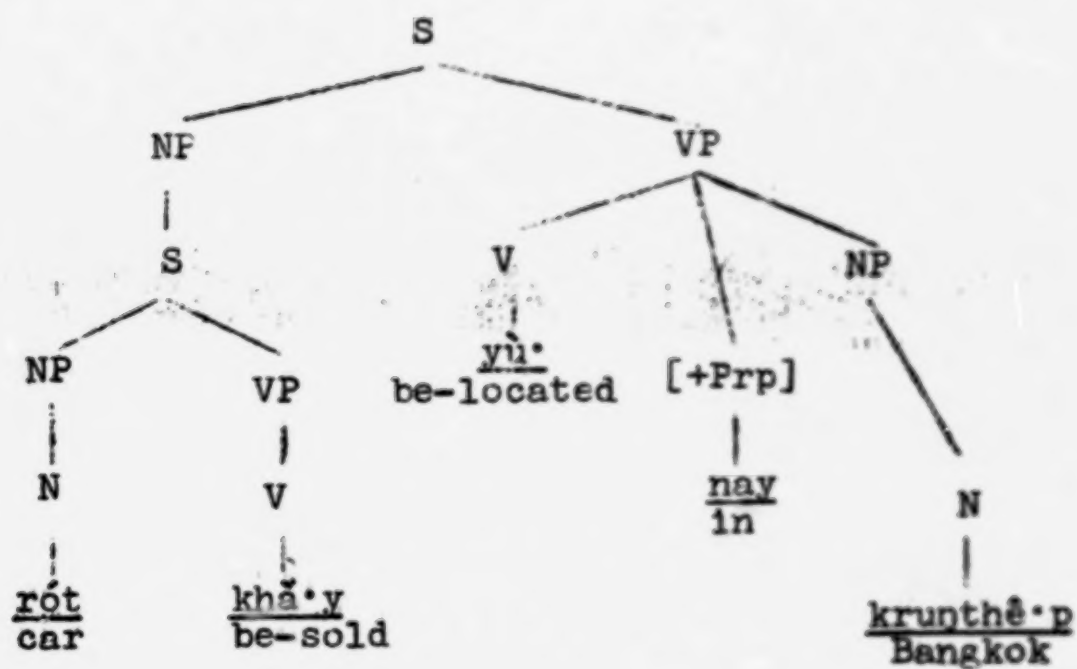
4.2 The Relative Pronoun Insert Transformation.

T Rel Inst	<N ₁	X	<Z		<N ₁ >NP	W>S	Y>NP
	1	2	3		4	5	6
obl→	1	2	3	<[+Pron]	4>NP	5	6
				[+Rel]			

This rule inserts a relative pronoun into a noun phrase in an embedded sentence if the embedded noun phrase consists of a noun which matches in reference the head noun of the embedding noun phrase, and if the embedded sentence is embedded in a noun phrase. If T Rel Inst were allowed to apply before T Amb Vb, the latter would have to be complicated by the optional presence of a relative pronoun as an immediate constituent of the object noun of the ambivalent verb, plus the



a.



b.

Figure 6. -- Operation of the Ambivalent Verb Transformation.

- a. Input.
b. Output.

specification that the relative pronoun is preposed with the object noun, if present. By allowing the relative pronoun to be inserted after the operation of T Amb Vb, T Amb Vb applies in the same form whether or not the sentence is embedded in another sentence. Figure 7 illustrates T Rel Inst.

4.3 The Redundant Noun Deletion Transformation.

T Red N Del $\langle X (N_1) Y \langle W \langle \begin{bmatrix} +\text{Pron} \\ +\text{Rel} \end{bmatrix} N_1 \rangle_{NP} Z \rangle_S U (N_1) E \rangle_S$

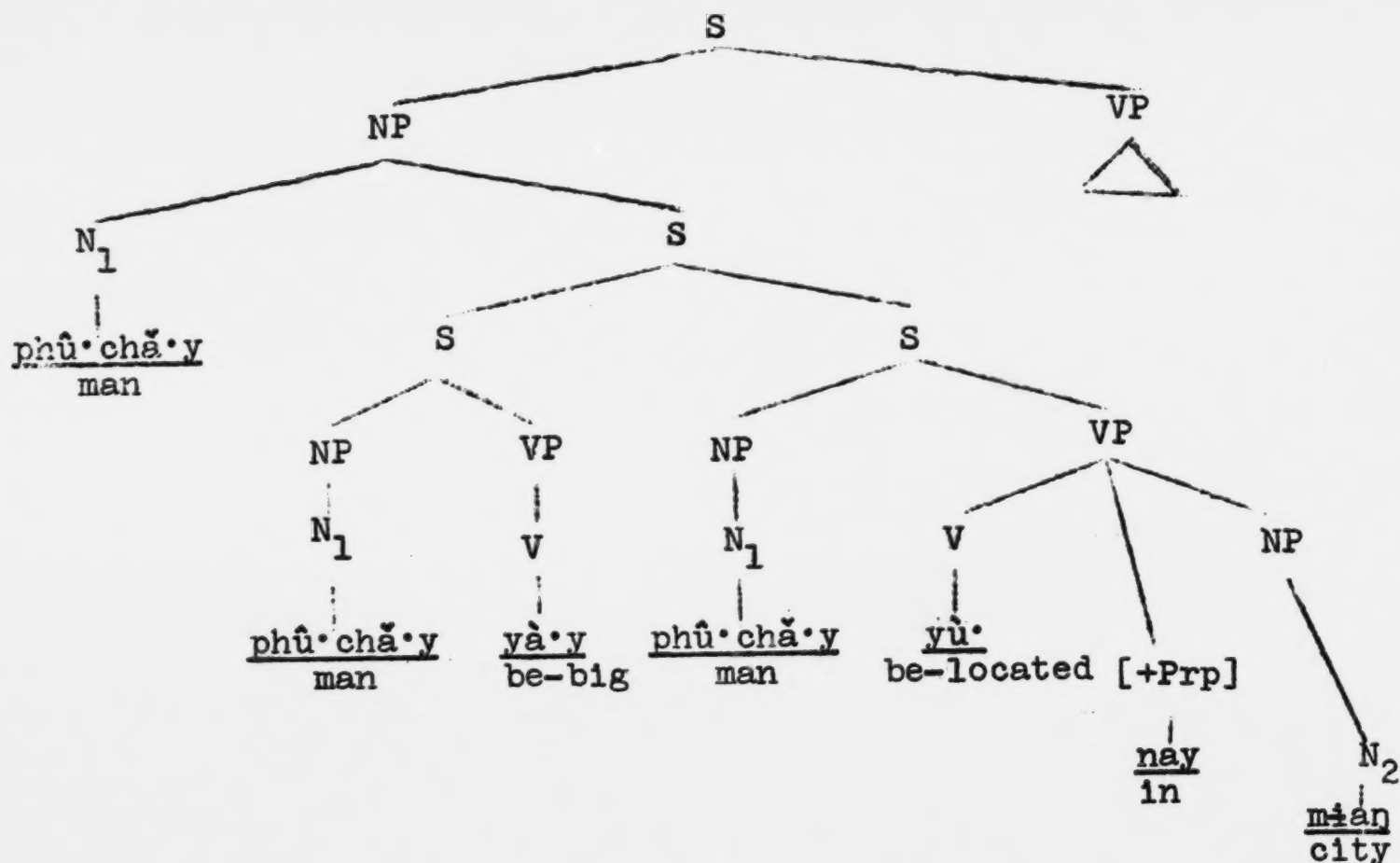
1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	∅	7	8	9	10

where 2 and 9 are not both absent.

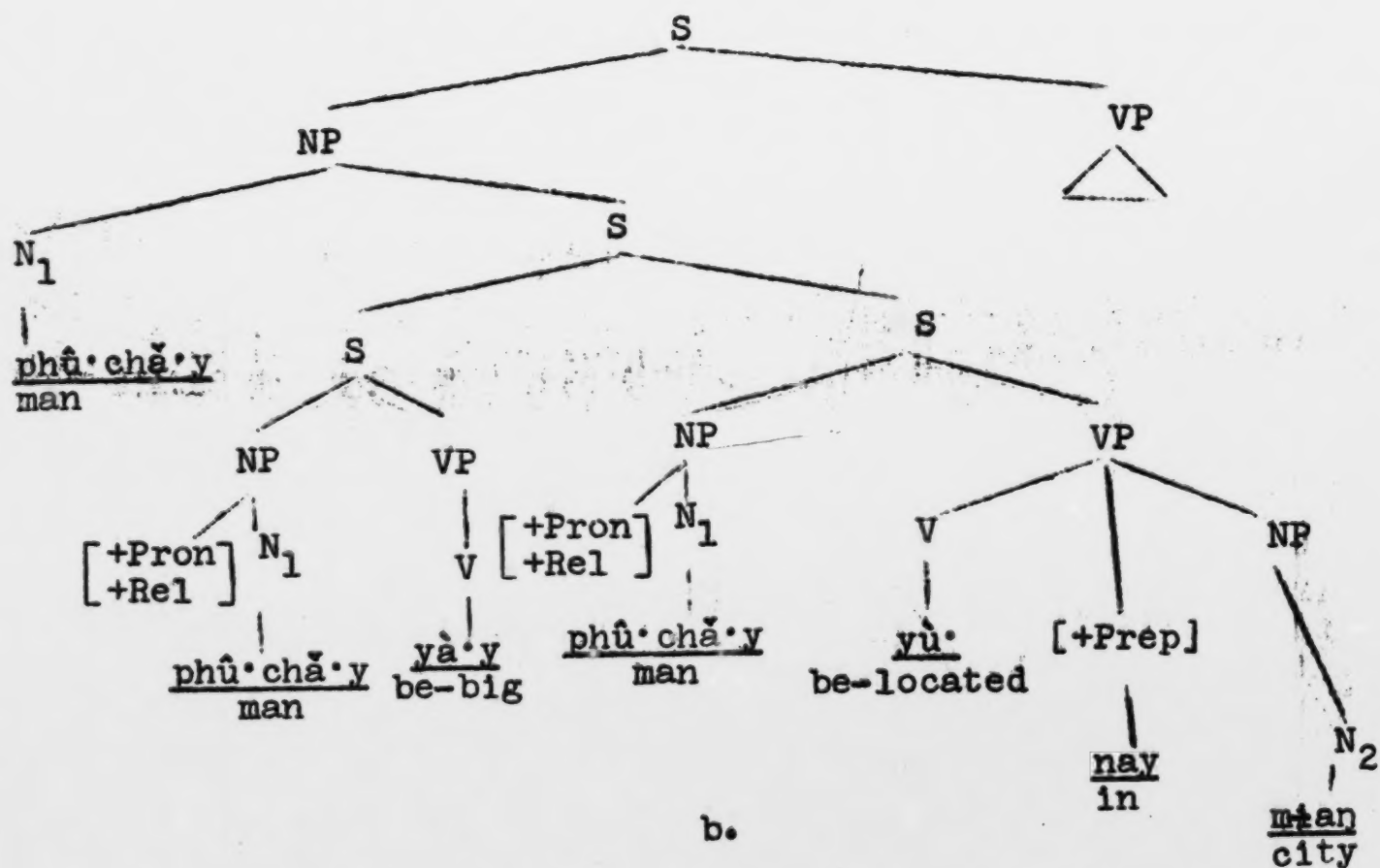
T Red N Del is designed to delete a noun in a noun phrase in which it and, possibly, a relative pronoun are the only constituents, if it is identical in reference with another noun in the embedding sentence. This rule is equivalent to Rosenbaum's identity erasure transformation.⁷ Our rule differs from his in a number of ways. Most crucially, it operates on nouns rather than noun phrases. In an English structure, for example, which can be represented as follows:

$\langle \text{The fat boy in the window} \langle \text{the (fat) boy (in the window)} \text{ is my brother} \rangle_S \rangle_{NP}$

the items in parenthesis are not absolutely essential for the generation of the noun phrase "The fat boy in the window who is my brother". What is essential is that the boy referred to in the embedded sentence be identical with the fat boy in the window in the embedding noun phrase. If T Red N Del is set up so as to delete only a noun, the noun phrase can remain to dominate a relative pronoun, thus simplifying the following



a.



b.

Figure 7. — Operation of the Relative Pronoun Insert Transformation

a. Input.

b. Output.

rule. Rosenbaum also requires that the erasing noun be positively identified. If there are two noun phrases which are identical to an embedded noun phrase, it is the closer one which causes the erasure. We only require that there be at least one identical noun in the embedding structure and do not require positive identification of the erasing noun. Apparently, there are no problems in our grammar caused by allowing this ambiguity. Finally, Rosenbaum requires a complementizing morpheme (such as the 'that' of the sentence "I doubt that they will come".) in the structural index of his transformation. This would be inappropriate in our grammar, since T Red N Del is designed primarily to handle relative clauses and not complements. T Rel N Del is illustrated in Figure 8.

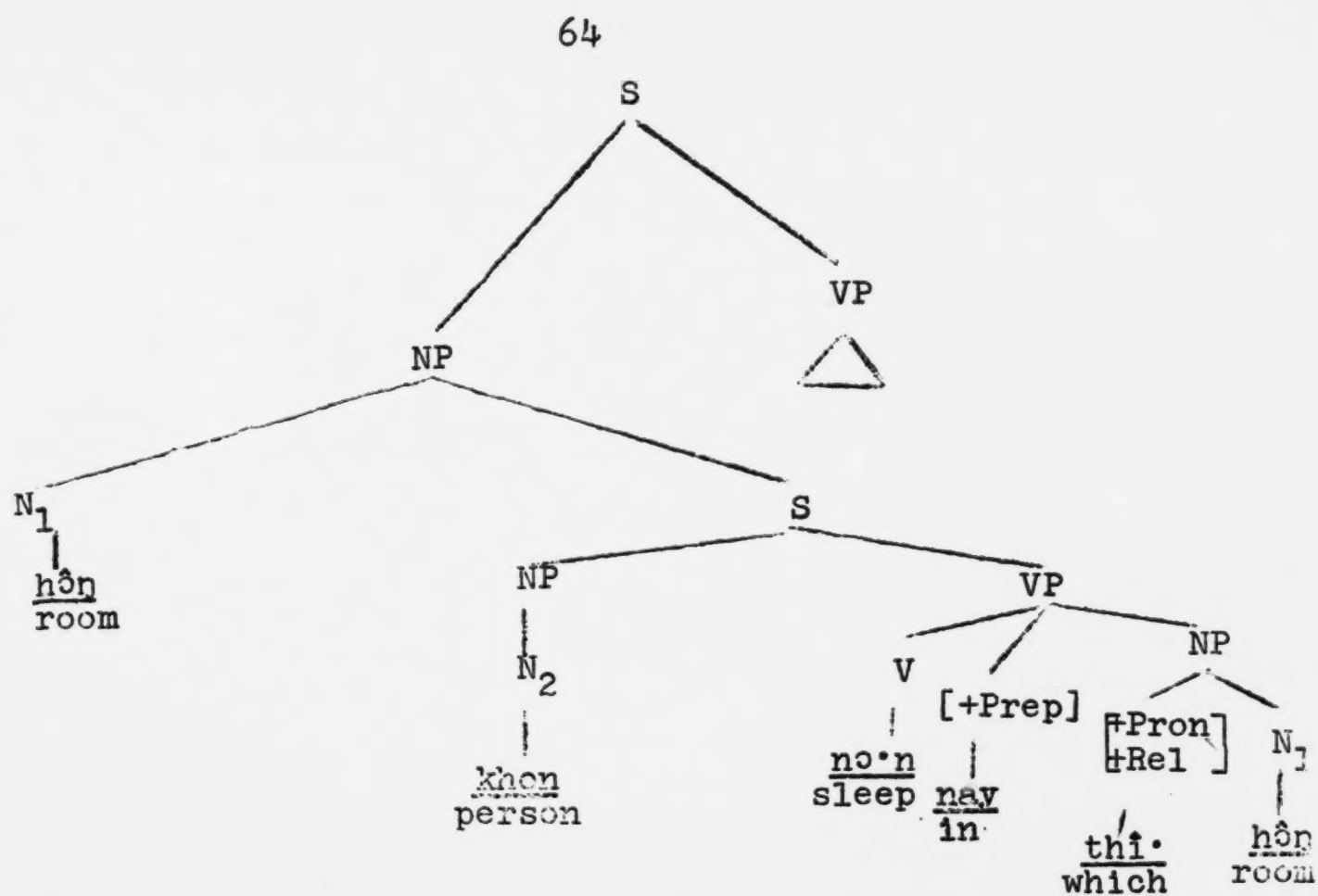
4.4 The Relative Pronoun Pre-position Transformation.

T Rel Pro Pre <X <Z [+Pron] W>S Y>S
 [+Rel]
 obl 1 2 3 4 5
 1 3 2 4 5

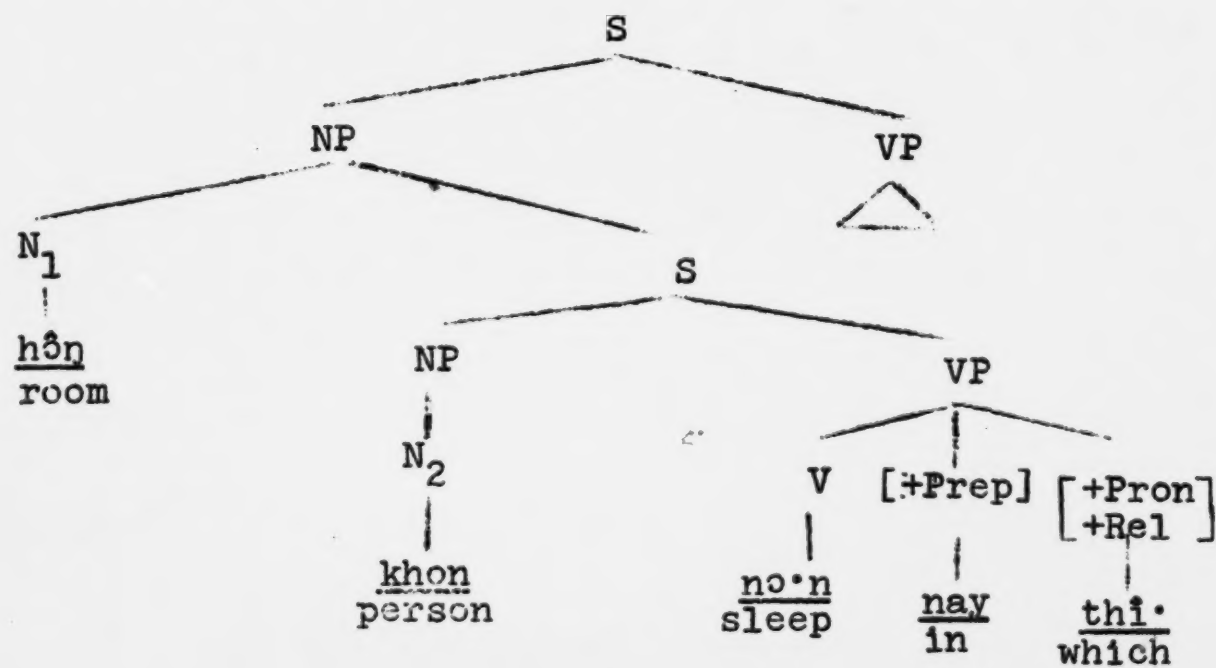
This rule takes a relative pronoun and places it at the front of the sentence in which it is a constituent. As in English, this rule applies vacuously when the subject has been relativized (and 2 is therefore null). Since this rule applies after T Red N Del, only the relative pronoun, and not the relative pronoun and its noun, is preposed. See Figure 9.

4.5 The Indirect Object Transformation.

T IO <NP <V Prp NP NP>VP X>S
obl → 1 2 3 4 5
1 2 4 3 5



a.

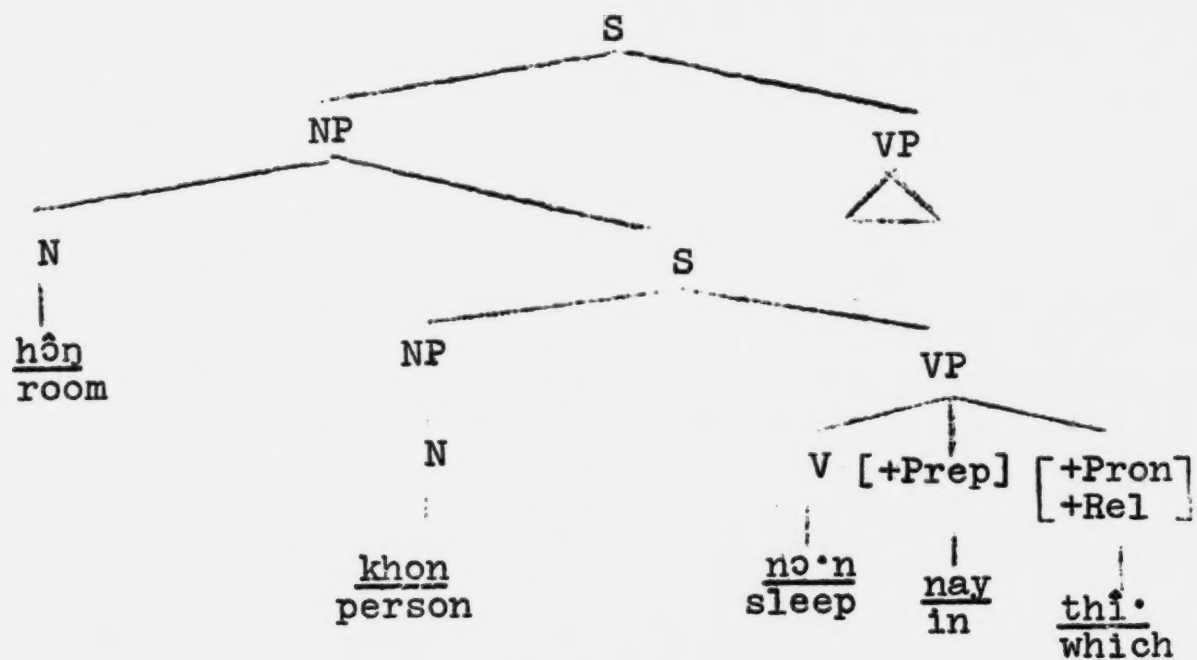


b.

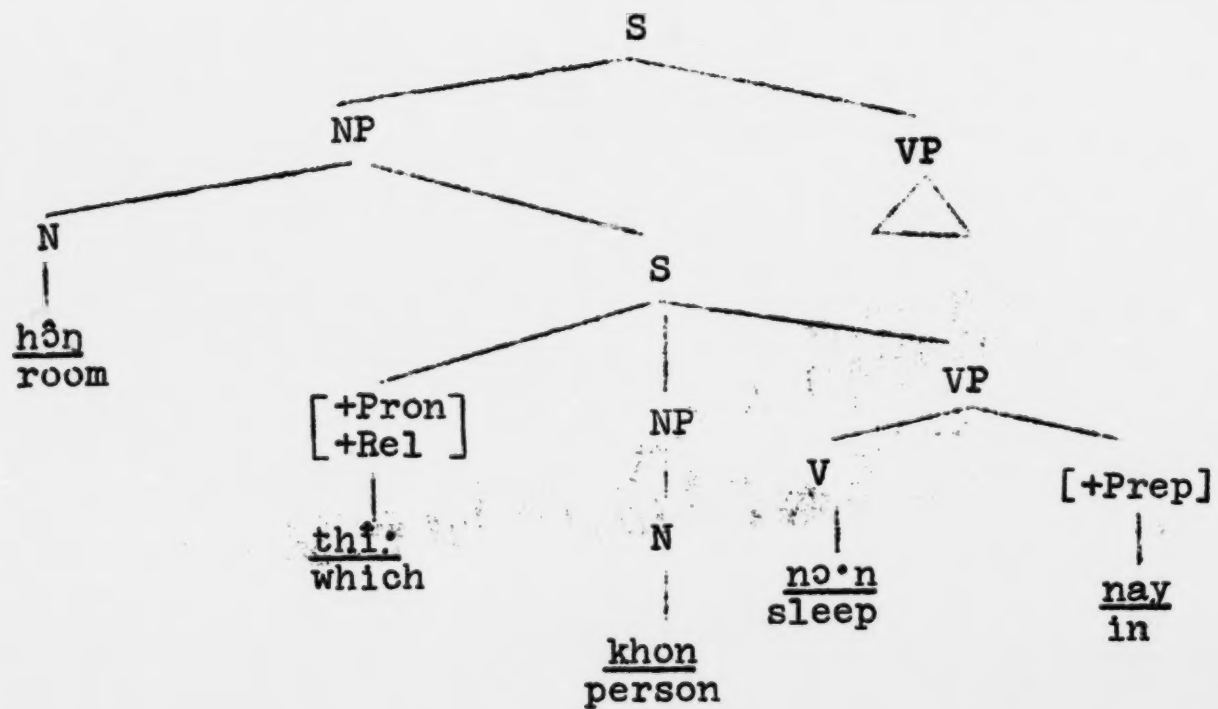
Figure 8. -- Operation of the Redundant Noun Deletion Transformation.

a. Input.

b. Output.



a.



b.

Figure 9. -- Operation of the Relative Pronoun Pre-position Transformation.

- a. Input.
- b. Output.

The Indirect Object Transformation places the direct object in a sentence which also contains an indirect object directly after the verb and before the preposition. This rule adjusts structures like:

khăw hây kê phôm nansăi
he give to I book

to the grammatical:

khăw hây nansăi kê phôm
he give book to I

"He gave the book to me."

More interesting to us will be the application of this rule to sentences containing the words chây sămràp 'to use for'.

T IO converts structures like:

<khon chây sămràp <khon rian tamra> NP tamra> S
person use for person study text text

to:

<khon chây tamra sămràp <khon rian tamra> NP> S
person use text for person study text

This sentence ultimately will become:

khon chây tamra sămràp (ka.n) rian
person use text for activity-study

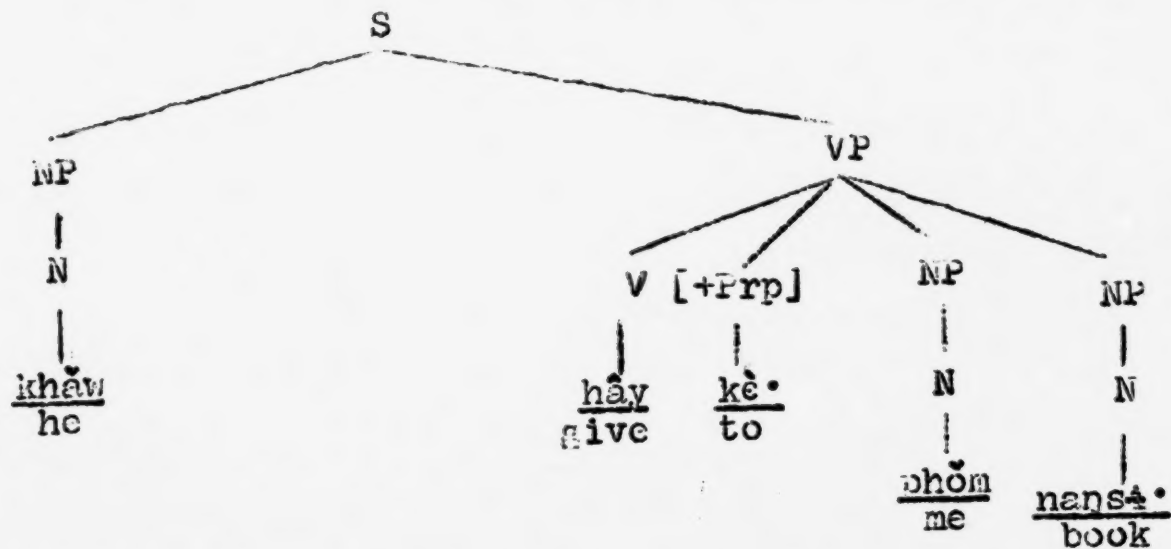
"People use the text for studying."⁸

Figure 10 illustrates this rule.

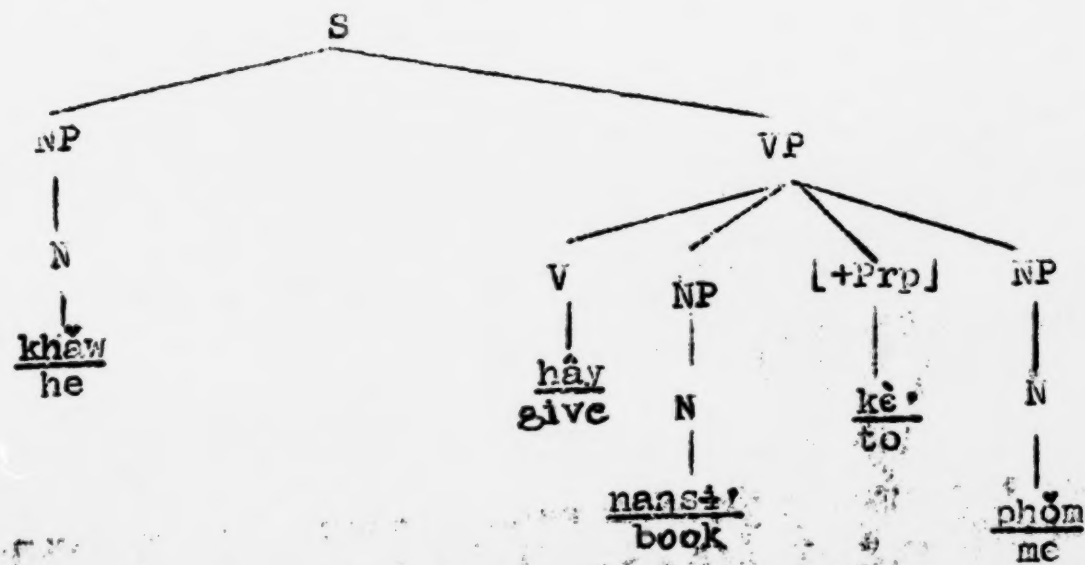
4.6 The chây Deletion Transformation.

T chây Del [+Pron] chây sămràp NP
[+Rel]
→ 1 2 3 4
1 ∅ 3 4

T chây Del is the first of two transformations which are simpler to state if they precede the deletion of the



a.



b.

Figure 10. -- Operation of the Indirect Object Transformation.

a. Input.

b. Output.

relative pronoun. Since the relative pronoun is inserted only in embedded sentences, no statement of the embedded nature of the sentence involved need be stated in the structural index if the relative pronoun is still present. Notice that the cháy 'to use' is only deleted when it appears directly before the preposition sămráp 'for'. But T IO, which immediately precedes this rule, places the direct object between verbs like cháy and prepositions like sămráp. If T IO has applied in this way, T cháy Del is blocked. But cháy in Thai may only be deleted if it means, not 'to use', but 'to be used'; that is, if the ambivalent verb transformation has applied. The ambivalent verb transformation, which applies before the indirect object transformation, preposes the object to subject position. When it applies, the indirect object transformation cannot apply and the sentence structure meets the structural requirements for T cháy Del. If these three rules apply in this order, the correct structures are most efficiently generated. In the overwhelming majority of sentences which were not deliberately solicited with this structure in view, cháy actually was deleted under the conditions specified in T cháy Del. However, Mr. Sirésamphan would accept as grammatical, sentences in which cháy is retained. A comparison of the following two sets of sentences, the first of which contain cháy as a main verb and the second of which contain instances of cháy deletion, illustrate the operation of this rule.

cháy as main verb:

mī·t lēm nán cháy săm·ráp hàn nĩa⁹
knife (clf.) that be-used for cut-up meat

"That knife is used for cutting up meat."

khăw cháy khwa·y săm·ráp thây na· lế? lâ·k kwian¹⁰
they use water-buffalo for plow field and pull cart

"They use water buffaloes for plowing fields and pulling carts."

cháy deleted:

khăw tham sáp·han săm·ráp khă·m mế·ná·m¹¹
they make bridge for cross river

"They make bridges for crossing rivers."

hế· săm·ráp thô·t pla· tham dây ?aray¹²
net for catch fish be-made with what

"What are nets for catching fish made of?"

It is clear that thi· cháy "which are used" has been deleted from each of the second set of sentences.

4.7 The yù· Deletion Transformation.

T yù· Del $\langle \langle \{ \begin{matrix} S \\ [+Pron] \\ [+Rel] \end{matrix} \} \rangle_{NP} \langle \underline{yù·} \text{ Prp NP} \rangle_{VP} \rangle_S$

1	2	3	4
→ 1	∅	3	4

Optional if 1 is $\begin{bmatrix} +Pron \\ +Rel \end{bmatrix}$, obligatory if 1 is S.

If the verb yù· 'to be located' is the main verb of a relative clause, it may be deleted and if it is the main verb of a sentence whose subject is a sentence, it must be deleted. The first half of the rule prepares the way for the generation of sentences with locative adverbials. The structure:

$\langle \langle \underline{khăw} \text{ tham} \underline{na·n} \rangle_{NP} \langle \underline{yù·} \text{ nay } \underline{mian} \rangle_{VP} \rangle_S$
he work be-located in city

becomes, by T yù Del.:

$\langle \langle \text{khăw} \text{ thamna} \cdot \text{n} \rangle_{\text{NP}} \langle \text{nay} \text{ mian} \rangle_{\text{VP}} \rangle_{\text{S}}$
 he work in city

and ultimately emerges from the grammar as:

khăw thamna·n nay mian
 he work in city
 "He works in the city."

Similarly, the noun phrase:

khon nay mian
 person in city
 "People in cities"

originates, via T yù Del, from the structure underlying:

khon thi· yù· nay mian
 person who be-located in city
 "People who are in cities"

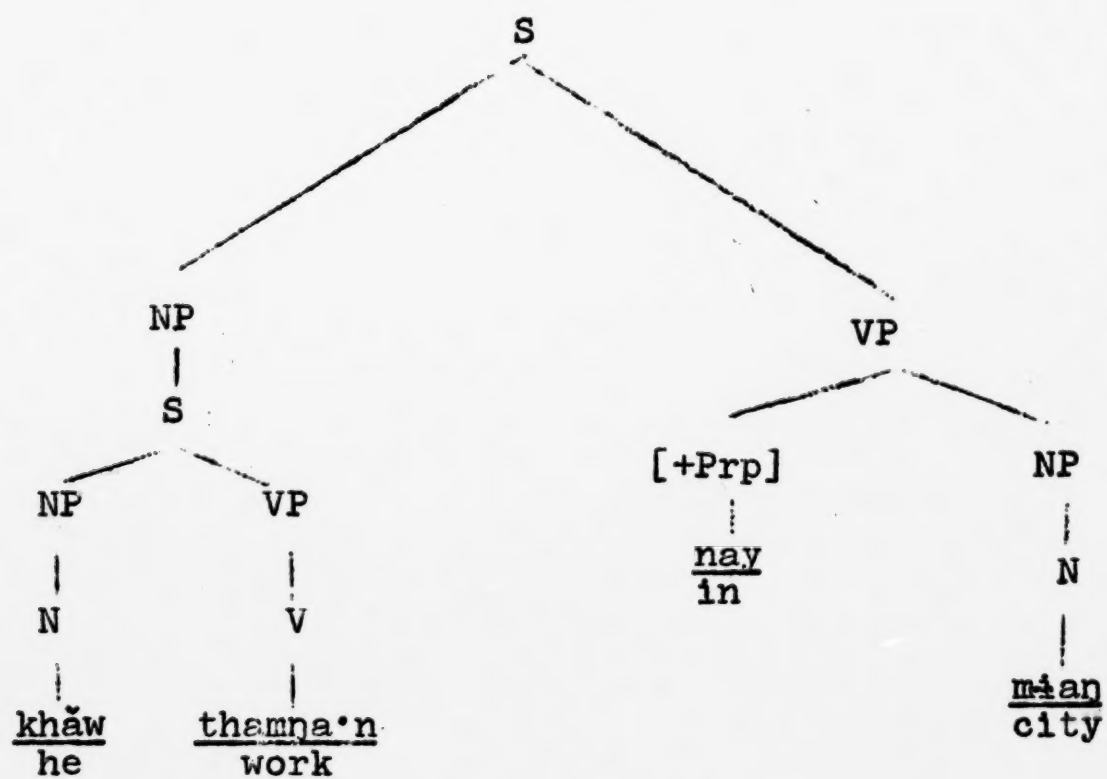
Like T cháy Del, it is easier to state this rule if it is allowed to apply before the relative pronoun is deleted.

4.8 The Adverbial Promotion Transformation.

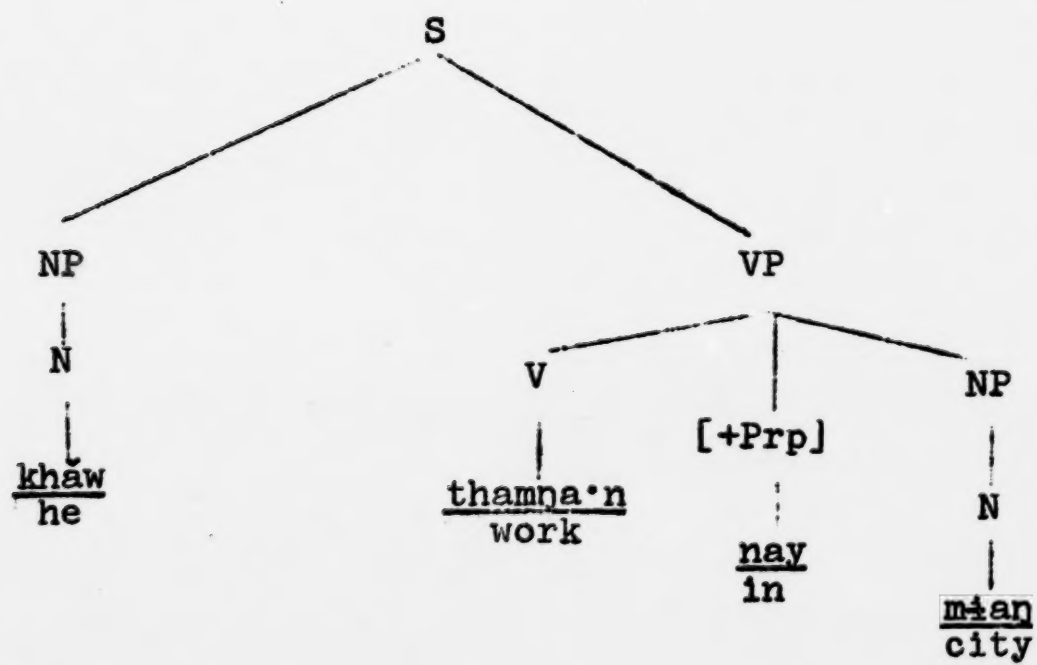
T Adv Prom $\langle \langle \langle \text{NP} \text{ VP} \rangle_{\text{S}} \rangle_{\text{NP}} \langle \text{Prp} \text{ NP} \rangle_{\text{VP}} \rangle_{\text{S}}$

obl₁ 1 2 3 4
 1 2>3+4 ∅ ∅

The adverbial promotion transformation takes the prepositional phrases left by the yù deletion transformation (and perhaps from other sources as well) and places them under the verb phrase of the sentence subject, thus creating the surface structure of such adverbials.¹³ This rule is not strictly ordered as the others in this section are, but it must apply before the nominalization transformation. It has been placed here for ease of exposition. See Figure 11.



a.



b.

Figure 11. — Operation of the Adverbial Promotion Transformation.

a. Input.

b. Output.

4.9 The Pronoun Deletion Transformation.

T Pron Del	X	[+Pron]	Y
	1	2	3
→	1	∅	3

T Pron Del is a broad rule which simply deletes all pronouns, personal or relative, optionally in any environment whatsoever. This lack of restriction may seem to be far too broad, but examination of the facts of Thai seem to show that this is exactly what happens. In Spoken Thai, Haas and Subhanka give the following examples of the deletion of personal pronouns:¹⁴

khə·y sǎa na:n
wait lose long-time

"(I)'ve been waiting a long time."

nák wā· cà? mây mā· sǎa lǎ·w
think that will not come lose be-finished

"(I) thought (you) weren't coming."

thammay chá· nák lǎ? khǎp
why be-slow very (ptl.) sir

"How's come (you)'re so late?"

With the examples and a discussion for learners of Thai, the authors make the statement that "...pronoun omissions in general are much more common in Thai than in English."¹⁵

Udom Warotamasikkhadit, in his doctoral dissertation, has an unrestricted rule for the deletion of personal pronouns which is similar to T Pron Del.¹⁶

It is also reasonably clear that relative pronouns are as freely deleteable. The following examples containing clauses which stand in relative relationship to some noun in the sentence,

but lack the relative pronoun, illustrate the point.

khăw mák cả? khian bồ·k way wâ· khon thĩ
they likely will write tell keep that person hold

nansỹ· pen phê·t ?áray
book be sex what

"They usually write down what sex the person who holds the passbook is."

mi· khonkhrua pen khon cátka·n ?áhă·n
have cook be person prepare food

"We have a cook who is a person who prepares food."

khăw mi· sĩa lé? ka·nke·nchánnay thúk chánít
they have undershirt and undershorts all kind

khă·y thĩ· hă·n nán
sell at store that

"They have all sorts of undershirts and undershorts which are sold at that store."

hố·n nĩ· mi· prātu·lèk pít wáy yanněnnă·
room this have iron-door shut be-secure heavily

"This room had an iron door which was heavily locked."¹⁷

Some of the above examples (e.g. the first one, if re-translated "...the person holding the (pass)book...") suggest a rule in English which also deletes the relative pronoun. However, in English the rule only applies when the pronoun is followed by 'is'. In his discussion of this point, Lees cites a series of six examples of widely varying provenience:

The man (who is) standing there is John.
The man (who is) taken there is John.
The man (who is) to go there is John.
The man (who is) over there is John.
The man (who is) for us is John.
The man (who is) asleep there is John.¹⁸

Of the above examples, the fifth has no direct equivalent in Thai. Of the others, none involves anything corresponding to 'is' in English. The first sentence in Thai would have the

progressive-aspect morpheme, kamlan. The second would involve an ambivalent verb or a more paraphrastic construction not involving any word for 'is'. The third would involve a separate modal. The fourth example would be translated using the word yù 'to be located'. The last example requires the use of a verb meaning 'to sleep' and no word for 'is'. So it is clear that no such restriction applies in Thai. Nor does any other type of restriction seem necessary. It appears that the Thai speaker is free to delete any pronoun any time he feels that no ambiguity will result.

Once established, T Pron Del is seen to be crucial in the derivation of several other constructions. The nominal use of adjectival verbs is a direct result of the application of this transformation. The sentence dek yà·y pay klàp bâ·n "The big boy is going home." is derived by this rule from dek thi· yà·y pay klàp bâ·n "The boy who is big is going home". This construction directly underlies noun compounds like hō·nwa·n 'vacant room'. T Pron Del, as we shall see, figures in the derivation of most other noun compounds as well. If it applies before the apposition and possessive transformations, these rules are simplified by being relieved of the necessity of specifying the subjects of certain embedded sentences. Similarly, if these rules were allowed to apply first, T Pron Del would be complicated because it would have to be written so as to apply obligatorily in the environments generated by those rules.

4.10 The Apposition Transformation.

T App <N <pen <N X>_{NP}>_S>_{NP} 1≠3
 → 1 2 3 4
 1 ∅ 3 4

The apposition rule deletes pen 'is' from sentences embedded in noun phrases, leaving the following noun phrase in apposition. The two nouns involved cannot be the same to avoid ungrammatical structures like *phũ·chǎ·y, phũ·chǎ·y... "The man, the man...". Note that T Pron Del has deleted the relativized subject of pen.

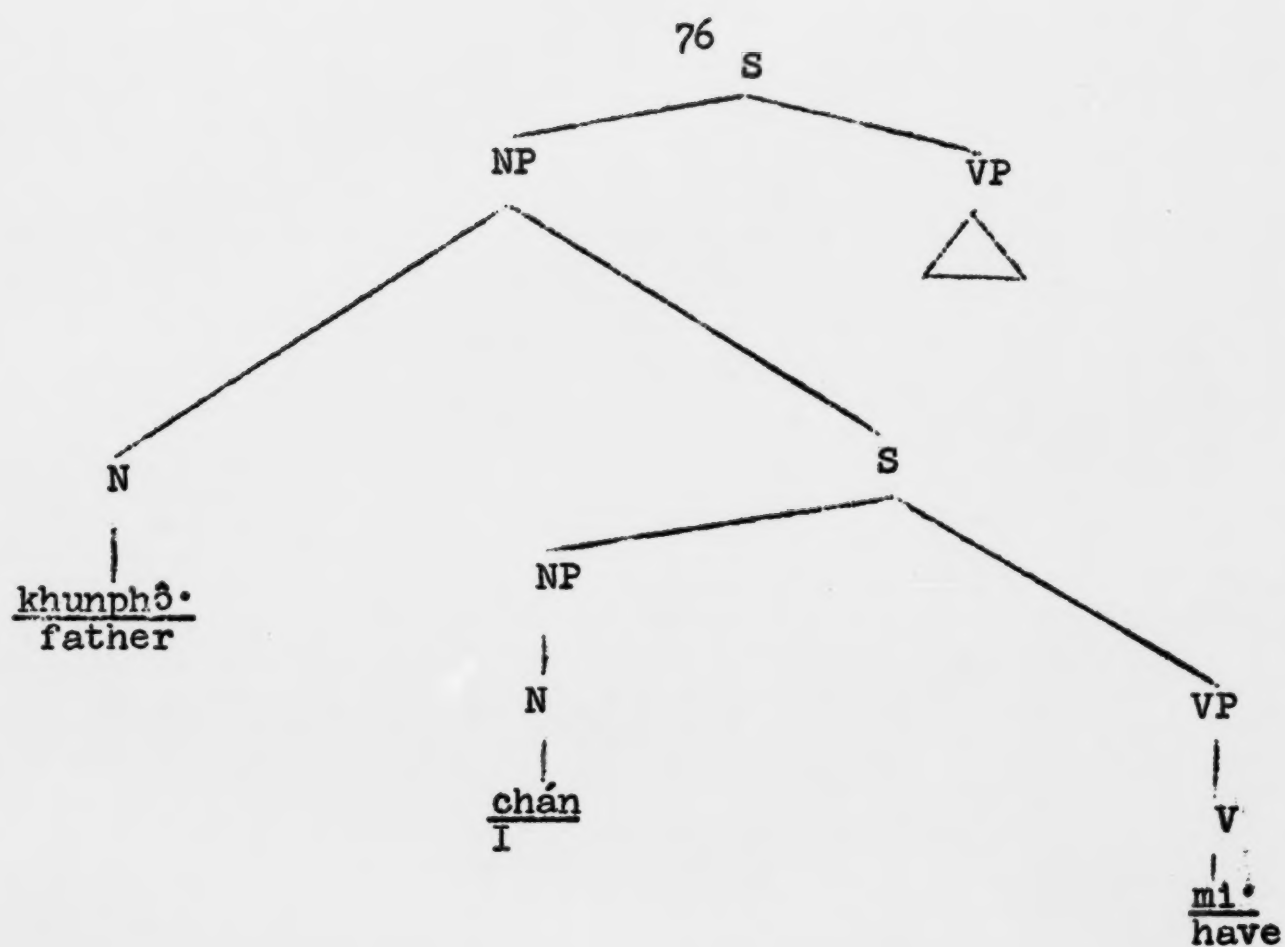
4.11 The Possessive Transformation.

T Poss <NP X <NP mi·>_S>_{NP}
 → 1 2 3 4
 1 2 khǒ·ŋ ∅

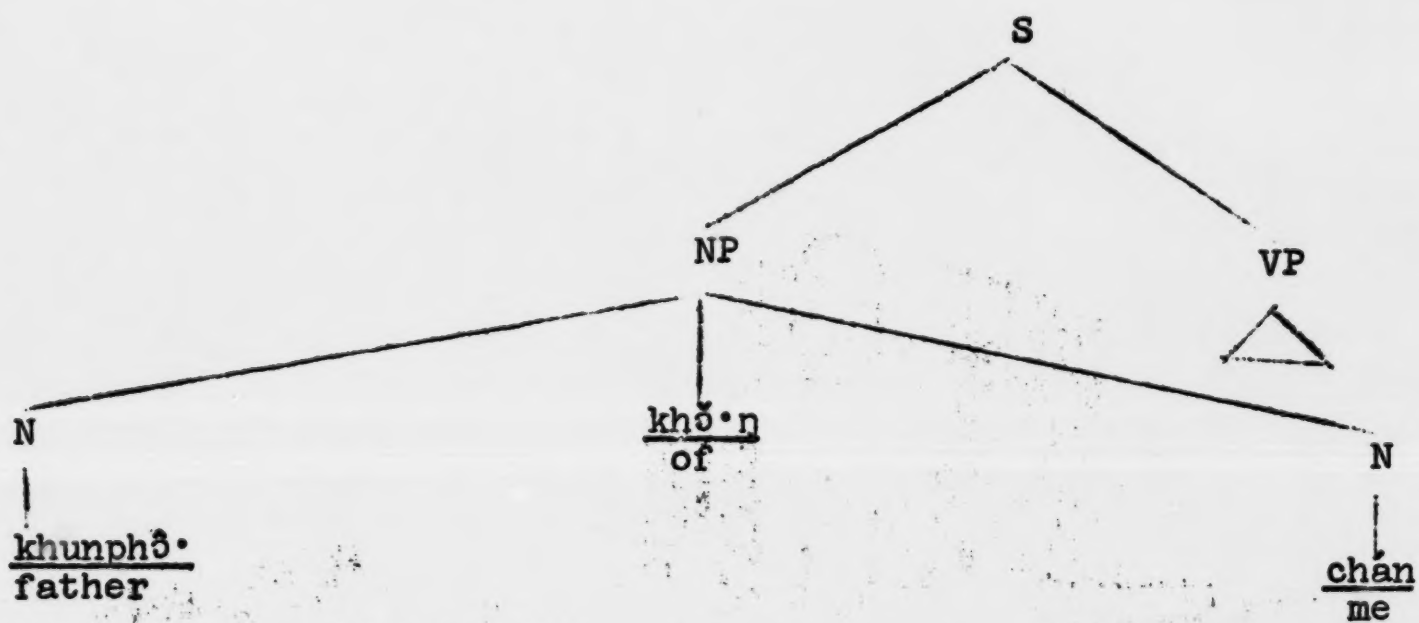
The possessive rule takes relative clauses of which the main verb is mi· 'to have', deletes mi·, and inserts khǒ·ŋ 'of'. The object of mi·, which is identical in reference to the head noun of the noun phrase marked 1 in the rule, has been relativized and deleted, and its relative pronoun has been pre-posed before application of T Poss. Hence, the order of T Poss is important with respect to T Rel Inst, T Red N Del, and T Rel Pro Pre; it must follow them. Application of the possessive transformation derives khunphô· khǒ·ŋ chàn 'my father' (lit. 'the father of me') from khunphô· chàn mi· 'the father I have'. Figure 12 illustrates T Poss.

4.12 The khǒ·ŋ Deletion Transformation.

T khǒ·ŋ Del <NP X khǒ·ŋ NP>_{NP}
 → 1 2 3 4
 1 2 ∅ 4



a.



b.

Figure 12. -- Operation of the Possessive Transformation.

- a. Input.
- b. Output.

The khǒ·n deletion rule optionally deletes the word khǒ·n 'of' generated by the possessive transformation.

T khǒ·n Del, if applied, converts khunphô· khǒ·n chán 'the father of me' to khunphô· chán 'my father'.

4.13 The Nominalization Transformation.

T Nom $\left[\begin{array}{c} +N \\ +Nom \end{array} \right] \langle NP \ VP \ X \rangle_S \rangle_{NP}$
 $\xrightarrow{obl} \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 1 > 3 & \emptyset & \emptyset & 4 \end{array} \text{ khǒ·n } 2$

If a nominalizing noun occurs immediately before a sentence embedded in a noun phrase, the verb phrase of the embedded sentence is inserted under the noun phrase node and directly to the right of the nominalizing noun. At the same time, the subject of the embedded sentence, preceded by khǒ·n 'of', is transposed to the end of the embedded sentence. The nominalizing nouns in Thai are ka·n 'activity' and khwa·m 'essence'. Selection of the first results in what Lees has called Action nominals and selection of the second yields Abstractive nominals.¹⁹ The difference between the two are neatly illustrated in an example of Mary Haas.²⁰ ka·n with the verb fǎn 'to dream' gives ka·nfǎn 'dreaming'. With khwa·m, the result is khwa·mfǎn 'a dream'. In other words, ka·nfǎn refers to the activity and khwa·mfǎn to the quality of the concept 'to dream'. We call ka·n and khwa·m nominalizing nouns because they are capable of being used as independent nouns. It is this fact that we can exploit by means of the proposed derivation. The input to T Nom is derived by the application

of pronoun deletion and apposition to the structure underlying the following strings:

ka·n (or khwa·m) thi· pen S
activity (or essence) which is

where the 'S' stands for a sentence of which the verb phrase is to be nominalized.

As in English, the subject of the nominalized verb is transposed to the end of the interior sentence with the insertion of a word meaning 'of'. Unlike English, only the subject is so treated. As a result, the Thai phrase corresponding to "the killing of the king" is unambiguous. With khǒ·n 'of', the phrase can mean only that the king did the killing. The two Thai forms are the following:

ka·nkhâ· khǒ·n pràmähä·kàsà·t
activity-kill of king

"The killing of the king" (the king kills someone)

ka·nkhâ· pràmähä·kàsà·t
activity-kill king

"The killing of the king" (someone kills the king)

Application of T Nom in the derivation of the former is illustrated in Figure 13.²¹

A study of Figure 13 shows that the output structure meets the structural index of T khǒ·n Del (where the first noun phrase is ka·nkhâ·, X is null, and the second noun phrase is pràmähä·kàsà·t). But khǒ·n cannot be deleted if it stands between a nominalized verb and its subject. Hence it is crucial that khǒ·n-deletion operate after the possessive rule, but before khǒ·n is generated by the nominalization transformation.

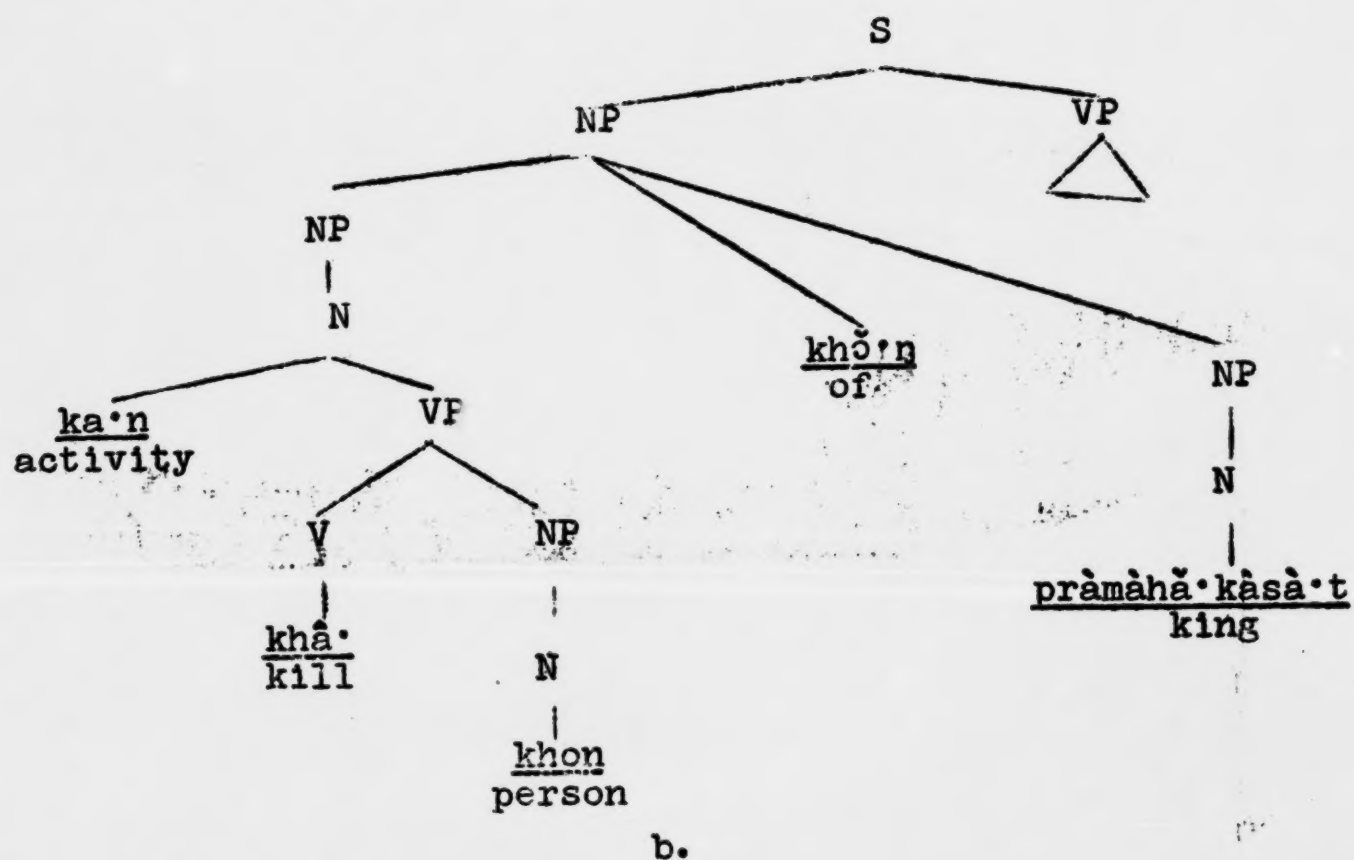
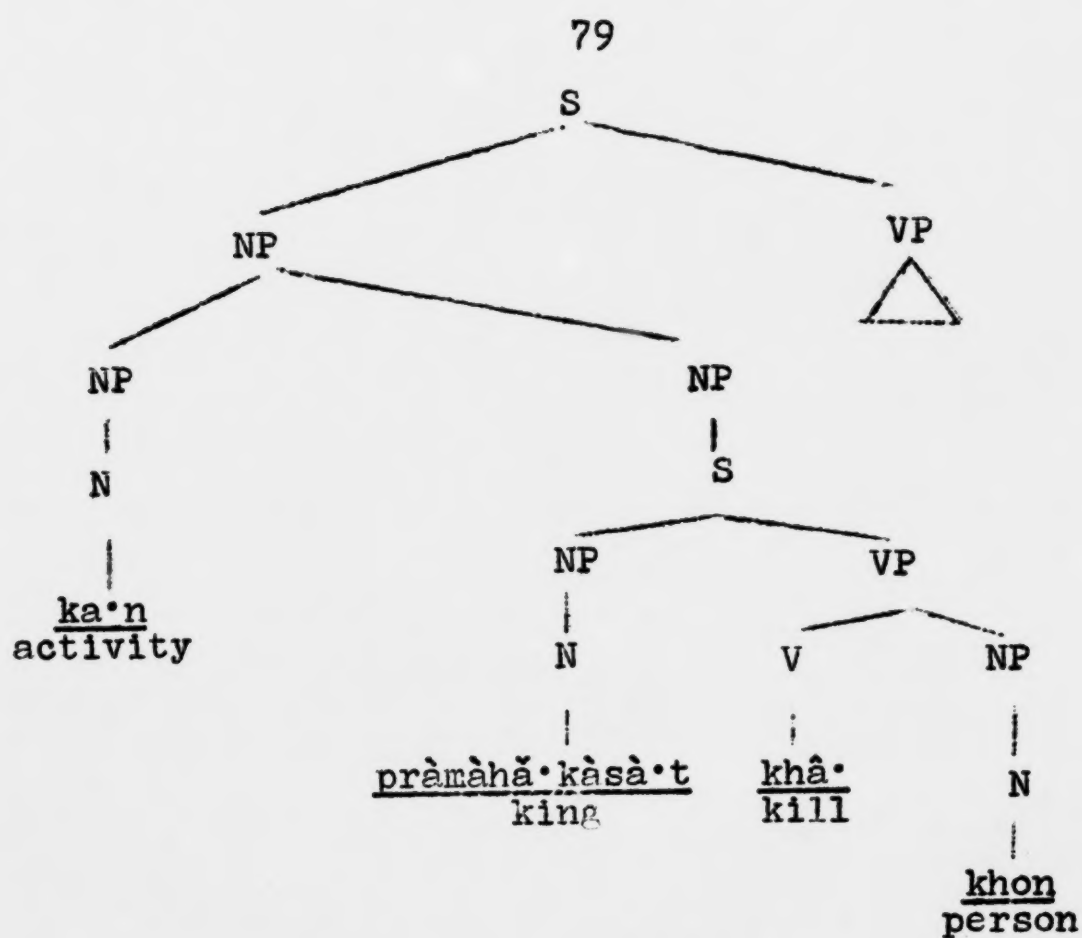


Figure 13. -- Operation of the Nominalization Transformation.

a. Input.

b. Output.

4.14 The Pro-form Deletion Transformation.

$$\begin{array}{cccccc}
 \text{T Pro Del } < \text{ka} \cdot \text{nV} & \text{X} & (\text{kh}\ddot{\text{o}} \cdot \text{n}) & \left[\begin{array}{c} +\text{N} \\ +\text{Pro} \end{array} \right] & \text{Y} >_{\text{NP}} \\
 \rightarrow & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & \emptyset & 5 \end{array}
 \end{array}$$

If either the subject or the object or both are Pro-forms, they may be deleted. The rule specifies that kh $\ddot{\text{o}}$ ·n may or may not be present. If it is, the following Pro-form is the original subject. If it is absent, the Pro-form must be the original object. The rule may be applied so as to delete Pro-forms in any combination. If both subject and object are Pro-forms, only the subject will be deleted if the rule is applied so as to interpret the object as "X", i.e. the number 2 item in the structural index. Similarly, only the object will be deleted if the sequence of kh $\ddot{\text{o}}$ ·n and subject Pro-form are interpreted as "Y", i.e. the number 5 item in the index. If both are present and to be deleted, the following procedure takes place:

$$\text{ka} \cdot \text{nV} \left[\begin{array}{c} +\text{N} \\ +\text{Pro} \end{array} \right] \text{kh}\ddot{\text{o}} \cdot \text{n} \left[\begin{array}{c} +\text{N} \\ +\text{Pro} \end{array} \right]$$

The rule interprets null as X (number 2) and kh $\ddot{\text{o}}$ ·n $\left[\begin{array}{c} +\text{N} \\ +\text{Pro} \end{array} \right]$ as Y (number 5). Thus the above structure meets the structural index of T Pro Del and the first Pro-form, the object, is deleted. This leaves the following structure:

$$\text{ka} \cdot \text{nV} \text{kh}\ddot{\text{o}} \cdot \text{n} \left[\begin{array}{c} +\text{N} \\ +\text{Pro} \end{array} \right]$$

This still meets the rule's requirements and it can be applied again, deleting $\left[\begin{array}{c} +\text{N} \\ +\text{Pro} \end{array} \right]$ and leaving ka·nV kh $\ddot{\text{o}}$ ·n. This falls

under the domain of the isolated preposition deletion rule, which deletes khǎ·n, leaving ka·nV.

It is likely that the deletion of Pro-forms is much more general than this rule allows, but this rule is adequate for this fragment of Thai grammar.

4.15 The Isolated Preposition Deletion Transformation.

T Is Prep Del <X Prp>_{NP}

$$\text{obl} \rightarrow \begin{matrix} 1 & 2 \\ 1 & \emptyset \end{matrix}$$

If a preposition is without an object because its object has been deleted by the redundant noun deletion or the Pro-form deletion transformations, it is obligatorily deleted. It may be noted that our relative pronoun pre-position transformation says nothing about prepositions. The preposition is left in its original position, like the common style of English as in "The man who you spoke with" and unlike elegant style, as in "The man with whom you spoke". But unlike either, the preposition is deleted altogether. Compare the following sentences:

khon tham kràdà·t nay ro·nna·n
person make paper in factory

"People make paper in a factory."

and ro·nna·n thi· khon tham kràdà·t
factory which person make paper

"A factory in which people make paper."

khon na·n nay hôn
person sleep in room

"A person sleeps in the room."

and hôn thi· khon no·n
room which person sleep

"A room in which a person sleeps."

As one might expect, Mr. Sirisamphan, the informant with whom I checked these sentences, indicated that these structures are uncommon, apparently since ambiguity is a likely result, but there is no other way to say a sentence in which the object of the preposition has been relativized. The operation of this rule will be abundantly illustrated in a variety of situations in chapter 4, since redundant noun deletion is a key rule in the derivation of a number of noun compounds. Its operation can be easily seen if one imagines that the preposition nay 'in' in Figure 9 were deleted.

4.16 The sāmrap Deletion Transformation.

T	<u>sāmrap</u>	Del	X	<u>sāmrap</u>	Y
			1	2	3
		→	1	∅	3

If cháy has been deleted, sāmrap 'for' may also be deleted. The following sentences were used to check the validity of this rule. Of each pair, the first comes from a published source and the second is a paraphrase supplied by Mr. Sirisamphan.²²

...tham prom pràdáp bā·n
make rug decorate house

...tham prom sāmrap pràdáp bā·n
make rug for decorate house

"...make rugs to decorate the house."

phôm yà·k cà? sá· nansá· sàk lêm pay ?à·n nay rót
I want will buy book just (clf.) go read on train

phôm yà·k cà? sá· nansá· sàk lêm sāmrap pay ?à·n
I want will buy book just (clf.) for go read

nay rót
on train

"I would like to buy just one book to read on the train."

4.17 The lé? Deletion Transformation.

T	<u>lé?</u>	Del	W	X	<u>lé?</u>	Y	Z
			1	2	3	4	5
		→	1	2	∅	4	5

Obligatory if 2 = 4.

lé? 'and' may be deleted anywhere and must be deleted if it conjoins two instances of the same item. This last provision is designed to set up structures immediately underlying reduplicated forms.

4.18 The ka·n Deletion Transformation.

T	<u>ka·n</u>	Del	X	< <u>ka·n</u>	V> _N	Y
			1	2	3	4
		→	1	∅	3	4

When ka·n 'activity' has undergone the nominalization transformation, it may be deleted. Its presence is apparently considered redundant in most contexts, and it is usually deleted. The cases in which it is not will prove crucial to the analysis of certain noun compounds.

The structures generable by these rules underlie our analysis of noun compounds in Thai.

Footnotes

¹Rosenbaum claims that the cyclic application principle has been called into question by recent research (Peter S. Rosenbaum, The Grammar of English Predicate Complement Constructions (Cambridge, Mass.: The MIT Press, 1967), p. ix.). I am not familiar with the arguments on the subject, and so I assume cyclic transformational application since it seems useful in dealing with our topic.

²Chomsky, Aspects, p. 138, 139.

³Jerrold J. Katz and Paul M. Postal, An Integrated Theory of Linguistic Descriptions (Cambridge, Mass.: The MIT Press, 1964), p. 80, 82.

⁴Stephen R. Anderson, "Concerning the Notion 'Base Component of a Transformational Grammar'" (Unpublished), 1967.

⁵D. Terence Langendoen, "The Accessibility of Deep (Semantic) Structures" (To appear in Roderick Jacobs and Peter S. Rosenbaum (eds.), Readings in English Transformational Grammar.)

⁶I have no explanation for the fact that the Thai sentence with the Progressive aspect marker kamlan undergoes verb promotion, although its structure is just like the completive sentence.

⁷Rosenbaum, op. cit., p. 6.

⁸This sentence, or the structure immediately underlying it, does not, however, appear in the derivation of tamra'rian 'study-text'. In order to derive tamra'rian, the ambivalent verb transformation must first apply, as we shall see immediately.

⁹Mary R. Haas and Heng R. Subhanka, Spoken Thai (New York: Henry Holt & Co., 1945), p. 268. Notice that T cháy Del does not apply if cháy is the main verb.

¹⁰Mary R. Haas, Thai Reader (Washington, D.C.: American Council of Learned Societies, 1954), p. xvi.

¹¹Haas and Subhanka, op. cit., p. 449.

¹²Haas, op. cit., p. lxxix.

¹³It seems reasonable to expect that a more comprehensive grammar of Thai would show us how we might generalize the structural index on T yü: Del so as to remove the disjunction and eliminate T Adv Prom altogether. It seems that it may well be, for example, that T Adv Prom is a special case of the verb phrase promotion rule discussed in chapter 1. Perhaps a larger grammar would show that the deletion of yü: follows naturally from the ordinary developments of promoted verb phrases. Note that Prp NP does not meet the requirements for VP, but we will assume it retains that label until after T Adv Prom.

¹⁴Haas and Subhanka, op. cit., p. 323, 324.

¹⁵Ibid.

¹⁶Udom Warotamasikkhadit, Thai Syntax: An Outline (Bangkok, Thailand: The College of Education Prasarnmitr, 1963), p. 54. However, he also states that deletion of the second person pronoun by this rule is the source of imperative sentences, a claim falsified by the second and third examples from Haas and Subhanka, above.

¹⁷The first three examples are from Haas and Subhanka, op. cit. The last is from my personally collected corpus.

¹⁸Robert B. Lees, The Grammar of English Nominalizations (Bloomington, Indiana: Indiana University Research Center in Anthropology, Folklore, and Linguistics, 1963), p. 91 - 94. The relevant rule is T58, p. 94.

¹⁹Cf. Lees, op. cit., p. 64 - 69, 85. It does not seem necessary to distinguish Action and Gerundive nominals in Thai.

²⁰Mary R. Haas, Thai-English Student's Dictionary (Stanford, Calif.: Stanford University Press, 1964), p. 82.

²¹The structure of the output of this transformation suggests an additional condition on the labeling of nodes. Any structure embedded under a category symbol (like N) should automatically lose its original structure and become a string of symbols only. This would prevent N in Figure 13 from dominating a verb phrase.

²²Haas, Thai Reader, op. cit., p. cxiii and Haas and Subhanka, op. cit., p. 327.

Chapter 4

Analysis of Compounds

1.0 Introduction. It will be the subject of this chapter to demonstrate how certain structures generated by the grammar proposed in chapters 2 and 3 can be converted into Thai noun compounds. We assume that most Thai noun compounds are derived from noun phrases in which sentences have been embedded as relative clauses. There seems to be considerable evidence for such an analysis. In the first place, there are a number of compounds in Thai which are composed of a collocation of words which are sentences in themselves and whose meanings reflect just these sentences. Two such words are:

<u>khonkhă.yphāk</u> 'vegetable vendor'	<u>dèklíankè?</u> 'shepherd boy'
<u>khon</u> 'person'	<u>dèk</u> 'child'
<u>khă.y</u> 'sell'	<u>lían</u> 'care for'
<u>phāk</u> 'vegetable'	<u>kè?</u> 'sheep'

In the second place, it is typically very easy for a Thai speaker to give sentences with relative clauses as definitions for compounds in his language when asked to. I have asked two informants to do this intensively with very good results. It appears less than artificial to propose that such sentences underlie a Thai speaker's competence in the use of noun compounds. Another piece of evidence for such an analysis involves a comparison with English. In English as Lees has shown, there are clear advantages to analyzing noun compounds as being derived from sentences with relative clauses.¹ Thus, the word

'blackboard' in English is derived from 'a board which is black' via the English noun modifier preposing rule to 'black board' and from this to the compound 'blackboard'. The corresponding Thai compound krada·ndam (krada·n 'board', dam 'to be black') has, in our analysis, a similar derivation from krada·n thi·dam 'a board which is black' to krada·n dam 'black board' to krada·ndam 'blackboard'. Since Thai does not have the noun modifier preposing rule which English has, both the noun phrase meaning 'black board' and the compound meaning 'blackboard' occur in the same constituent order which they have in the noun phrase with the relative clause. In an analysis which fails to relate the noun phrase to the compound, a contrastive statement between the two languages will have to include two statements about the order of noun modifiers where only one is really necessary.²

It is tacit in the analysis proposed above that one can distinguish between noun phrases and noun compounds, that is that one can tell a sequence of two separate words from the same two words compounded into one. In English, there is little problem since the stress pattern of 'black board' is different from that of 'blackboard'. In Thai, it is more of a problem. According to Richard Noss, junctural phenomena do the same work in Thai. Although he does not say this explicitly, Noss almost always writes compounds with a hyphen between the two members. The hyphen in Noss' phonological analysis means a rhythm phoneme representing "medium-short internal-syllable duration".³

The corresponding noun phrases are generally written with a space between the words, indicating "medium internal-syllable duration".⁴

I have not personally been able to distinguish such a contrast in my own work. However, this could be merely due to the fact that I do not know Thai sufficiently well to make the distinction. But apparently others who know Thai much better than I do are also unable to do so. William J. Gedney, the distinguished Thai scholar, considers making this distinction between compounds and collocations of words a difficult problem in working on Thai compounds.⁵ It appears that Udom Warotamasikkhadit does not make this distinction either, since he allows the same intonation with both me·w nō· 'stupid cat' and na·klāa 'salt marsh' (na· 'field', klāa 'salt').⁶

In her thesis on noun compounds, Miss Nisa Udomphol attempts to distinguish compounds from noun phrases on the basis of the behaviour of the members of putative compounds in certain syntactic frames.⁷ These frames are designed to test separability and vary according to the type of compound under discussion. For example, to distinguish compounds from collocations of two separate words when both members are nouns, Miss Udomphol selects the following three frames:⁸

thi·ni· mi· N₁ N₂ CL dīaw
here there-is single

N₂ CL ni· N₁ V_{1a} mā·k
this very

N₁ CL ni· pen khǒ·n N₂ CL nán
this be of that

In this notation, N stands for noun, CL for classifier, and V_{1a} for adjectival verb. Sequences of two separate words will occur in the two latter frames, but not in the former. As a result, khunce·tù· 'key to the cabinet' can be shown to be a sequence of two words since there is:

khunce· dô·k ní· pen khǒ·n tù· bay nán
key clf. this be of cabinet clf. single

"This key goes to that cabinet."

But no:

*thí·ní· mí· khunce·tù· bay díaw
here there-is cabinet-key clf. single

"Here is one cabinet key."⁹

On the other hand, hì·psían 'phonograph' (hì·p 'box', sían 'sound') is identifiable as a compound since there is no:

*hì·p bay ní· pen khǒ·n sían sían nán
box clf. this be of sound clf. that

"This box belongs to that sound."

although:

thí·ní· mí· hì·psían khrián díaw
here there-is phonograph clf. single

"Here is one phonograph."

But this criterion is only partly successful, as Miss Udomphol acknowledges. As it turns out, there are a number of words which fit in the frame which determines compounds and also one of the frames which determines sequences. For example:

thí·ní· mí· nám·na·w khùat díaw
here there-is lemon-juice jar(clf) single

"Here is one jar of lemon juice."

má·na·w lû·k ní· nám dí· má·k
lemon clf. this juice be-good much

"The juice of this lemon is very good."

So at the very best, this grammatical device is only partly capable of distinguishing compounds from collocations. Incidentally, on the basis of our analysis of Thai compounds, it is possible to predict just when the members of a compound will fit into a frame which separates them and when they will not. We shall limit our discussion to the first and third frames.¹⁰

As we shall show in this chapter, the basic mechanism in Thai for compound formation is a rule which takes certain material which is dominated by a Noun Phrase node and inserts it all under the head Noun of that Noun Phrase. Notice that this process does not affect the order or number of constituents in the least. If linear order of constituents is the only criterion, a given sequence of constituents will be both a noun phrase and a noun compound just in case the structure immediately preceding the noun compounding rule is a grammatical noun phrase. Even if it is, it still could fail to meet the frame test if it is a noun phrase of different type from that tested by the frame. For noun-noun compounds, there are two such structures, co-ordinate and apposition. Thus nám 'water, fluid' followed by kròt 'acid' could be either a compound meaning 'acid' or a noun phrase nám as head and kròt in apposition. This deficiency could easily be remedied by adding more test frames to the list. For the apposition case, we could add a frame like:

<u>thi.ni.</u>	<u>mi.</u>		N_1	<u>thi.</u>	<u>pen</u>	N_2
here	there-is			which	be	

If N_1 is nám and N_2 is kròt, this would mean "Here is fluid which is acid".

A number of noun compounds cannot fit into the prescribed frames because they have been assigned a sort of metaphorical meaning not derivable from the meaning of their members. For example, ta·nám 'underground waterway' (ta· 'eye', nám 'water') will not fit into any of these word-separating frames because of its meaning. But this is a problem which involves more than just compounds and thus is a whole separate topic.

A fourth situation in which frame tests can be invoked and results in a successful prediction by the frame test, is the case of the compound of which one member is a bound form. In this case, our rules will indeed generate a noun phrase which consists of the two separate nouns, but one of these has syntactic features which specify that the noun compounding rule must apply. The Thai word for 'bull elephant', chá·nphlá·y (chá·n 'elephant', phlá·y 'male' (used only with chá·n)), is such a word. The sequence chá·n followed by phlá·y would be a noun phrase with phlá·y in apposition to chá·n, except that phlá·y is so specified that the noun compounding rule must apply to such a noun phrase.

Finally, the frame technique will succeed in cases in which some constituents must be deleted after application of the noun compounding rule. The word kây·fá· 'pheasant' (kây 'chicken', fá· 'sky') will be compounded as *kây·bin·nay·fá· (kây 'chicken', bin 'to fly', nay 'in', fá· 'sky') with subsequent

deletion of bin and nay.

The frame technique for identifying single words which are made up of two other words is only partly successful and the reasons for its degree of success can be accounted for in a deep way by the kind of analysis under discussion in this chapter. With phonological and grammatical criteria of questionable value, the only solution seems to be the consensus of the speakers of Thai. As Miss Udomphol puts it:¹¹

The most serious limitation of this study is the lack of a definition of the term "word". No attempt has been made at such a definition since such an attempt would result in another research problem. The term is used here with its "common sense" definition.

2.0 Other Work on Thai Compounds.

2.1 Traditional Grammars in Thai. There are apparently a number of Thai grammarians who have had something to say about compounds in their own language. Miss Udomphol summarizes this work in her introductory chapter.¹² Of the four works she discusses, two of them seem rather naive, containing statements to the effect that the Thai people were forced to combine two words into compounds when they communicated with people who used polysyllabic languages and when they met new things in the south.¹³ As the other grammar book puts it, "It is because we know many new things and the simple words are not enough to use that we create new words from our simple words..."¹⁴ However, some insight into the nature of compounding seems to be characteristic of the work of Phya Upakitsilpasarn.¹⁵ He realizes that some compounds are reduced from a

phrase, which seems to hint at derivation from sentences. In addition, Upakitsilpasarn considers reduplication and combinations of learned words of Sanskrit-Pali origin to be of the same nature as compounding. We have taken the same position and learned words and reduplication will be discussed in chapter 5. According to Miss Udomphol, "the same ideas" are found in Reference Book of Thai Grammar for Secondary Education 4-5.¹⁶

2.2 Traditional Grammars in English. Thai noun compounds receive scant mention in traditional style grammars of Thai written in English. Campbell and Shaweevongse treat Thai noun compounds under three different headings.¹⁷ Those whose first members are words like pla 'fish' and bay 'leaf' are treated in the lesson on "Flaura and Fauna". Those whose heads are words like khon 'person', châ·n 'artisan', and cha·w 'inhabitant of' are treated in a lesson on "Personification of Nouns and Verbs" and "Nationalities". A number of others are discussed in a lesson on "Compound Words" which consists of a remark about Thai compound words being formed from a noun and one or more descriptive words, followed by a set of lists of compounds classified by their first members. The situation in Lanyon-Orgill is little better.¹⁸ Compounds are treated in the same section, but again this section consists of nothing but a terse description of the surface formation noun compounds followed by lists of compounds by first members.

2.3 Structural Grammars in English. Several American

linguists have written grammars of Thai in the structural style. Two of these, grammars by Haas and Subhanka and Anthony, French and Warotamasikkkhadit, are pedagogical grammars and as such are perhaps not to be expected to treat a topic like compound formation in great detail. Both of them describe the surface form of noun compounds together with comments about how they differ from English.¹⁹ In his Thai Reference Grammar, Richard Noss separates those compounds of which one member is bound from those in which both members occur independently.²⁰ The first set involves derivatives with prefixes and suffixes (i.e., the bound members). The second set is labeled "compounds" and here there is a description of the surface formation of compounds, but in somewhat more detail than in other grammars, and a few lists of examples. In the brief description of Thai in the introduction to the Thai-English Student's Dictionary, Mary E. Haas classifies Thai compounds according to the categories of their members.²¹ The four types listed are: 1) Noun + verb, 2) Noun + verb with object, 3) Noun + noun (head + attribute), and 4) Noun + noun (coordinate compound). This list is significant in two places. It is interesting that Dr. Haas calls the second type "Noun + verb with object" and not something like "Noun + verb + noun". This represents a tacit acknowledgment of structural relationships below the surface. Similarly, she distinguishes "Noun + noun (head + attribute)" from "Noun + noun (coordinate compound)" although there is nothing in the surface structure which would distinguish these.

2.4 Grammars by Thai Linguists in English. The most insightful work on compounds has been by Thai linguists who have been trained in linguistics in the West. Of these, two are particularly valuable. One is the above-mentioned thesis of Nisa Udomphol.²² The other is a doctoral dissertation by Udom Warotamasikkkhadit.²³ Miss Udomphol's work is done in structural linguistic theory and lists large numbers of compounds, drawn from the government Thai-Thai dictionary, according to the category and order of the members. However, the frames she uses to distinguish compounds which can only be compounds from sequences which can also be noun phrases generally reflect the structure of the sentences from which at least some of the compounds of a given type must be derived. As we have already seen, the frames used to identify noun-noun sequences which could be separate words reflect possession. As it turns out, a disproportionately large number of noun-noun compounds in Thai are of the possessive type. As we shall see, the compound khǒnkè? 'lamb's wool' (khǒn 'hair', kè? 'lamb'), is derived via khǒn khǒ·n kè? 'the hair of a lamb' from a noun phrase of the form:

khǒn thī· kè? mī·
hair which lamb have

"The hair which a lamb has"

Because this is the structure of so many of these compounds, it is not surprising that this frame is as successful as it is for identifying separable members.

For noun-verb-noun and noun-verb compounds, Miss Udomphol uses the frame:

N thĩ· V (N) CL nĩ· V_{1a} mā·k
 which this be-much

This structure represents a stage in the derivation of many such compounds. A noun compound of these shapes will fail to fit in this frame if one of the members is bound, if the meaning is metaphorical, or if some material is deleted after application of the compounding rule. A typical example of the last case is ro·nkhā·sāt 'slaughterhouse' (ro·n 'building', khā· 'to kill', sāt 'animal'). There is no *ro·n thĩ· khā· sāt ro·n nĩ·, so Miss Udomphol lists ro·nkhā·sāt as a sequence which could only be a compound. The reason for this is that the noun compounding rule creates *ro·nkhonkā·sāt (ro·n 'building', khon 'person', khā· 'to kill', sāt 'animal') ultimately from:²⁴

ro·n thĩ· khon khā· sāt
 building which person kill animal

"Building in which people kill animals"

After the compounding operation, khon is deleted.

Besides the noun-verb compounds discussed above, Miss Udomphol discusses separately noun-verb compounds of which the verb is an adjectival verb. Finding a frame to distinguish compounds from sequences of words was especially difficult here. For every noun-adjectival verb compound, there is a corresponding noun-adjectival verb noun phrase. As a result, the frame chosen by Miss Udomphol to identify compounds of this

type has nothing to do with grammar, but involves what is and what is not a contradiction. Both noun phrases and noun compounds of this type will occur in the following two frames:

$N + V_{1a(1)} \text{ CL } \frac{ní}{\text{this}} V_{1a(2)} \frac{mā·k}{\text{much}}$

$N + V_{1a(1)} \text{ CL } \frac{ní}{\text{this}} V_{1a(1)} \frac{mā·k}{\text{much}}$

A noun phrase can fill the first two slots of both frames.

Thus, using Miss Udomphol's example, we get:²⁵

$\frac{dèk}{\text{child}} \frac{sūay}{\text{be-beautiful}} \text{ Clf. } \frac{khon\ ní}{\text{this}} \frac{dī}{\text{be-good}} \frac{mā·k}{\text{be-much}}$

"This beautiful child is very good."

$\frac{dèk}{\text{child}} \frac{sūay}{\text{be-beautiful}} \text{ Clf. } \frac{khon\ ní}{\text{this}} \frac{sūay}{\text{be-beautiful}} \frac{mā·k}{\text{be-much}}$

"This beautiful child is very beautiful."

A noun compound, but not a noun phrase, will also fit into the following frame:

$N + V_{1a(1)} \text{ CL } \frac{ní}{\text{this}} \frac{mây}{\text{not}} V_{1a(1)} \frac{lə·y}{\text{at-all}}$

Thus the word námhǎ·m 'perfume' (nám 'fluid', hǎ·m 'to be fragrant') can appear in the sentence:

$\frac{námhǎ·m}{\text{perfume}} \frac{khuat}{\text{bottle}} \frac{ní}{\text{this}} \frac{mây}{\text{not}} \frac{hǎ·m}{\text{be-fragrant}} \frac{lə·y}{\text{at-all}}$

"This bottle of perfume (lit. fragrant fluid) isn't fragrant at all."

without contradiction. The same thing happens in English, where one can say "This blackboard isn't black at all" without contradiction, but "This thick board isn't thick at all" is self-contradictory. By the analysis which we are proposing, both the compound and the noun phrase receive the same derivation,

but the compound has the further process of compound-formation. One of two types of solutions is possible. Either one assumes that something happens to a noun compound after it is formed which accounts for this behaviour, or the derivation for compounds can be changed so that it is different from that of noun phrases. It would be possible, for instance, to add a phrase like 'is characterized by' to the derivation of noun compounds, but not to noun phrases. It is also conceivable that noun compounds which are frequently used get entered whole into the lexicon, where they may acquire other semantic features besides those contributed by the individual members. Without convincing evidence, we will assume that a solution of the latter type is appropriate, and that the syntactic derivation of the compound is parallel to that of the noun phrase up to compound formation itself. It seems that a solution of this type is necessary to handle compounds like mê.ná.m 'river' (mê 'mother', ná.m 'water') which retain very little of the meaning of original formation, but do retain at least a trace.

Dr. Warotamasikkhadit is the only linguist whose work I have seen who derives noun compounds from underlying sentences. Using an older model of transformational theory than that in this paper, he sets up a battery of generalized transformations which embed parts of one sentence into a noun phrase in another. In this way, sentences A and B are combined to form C:²⁶

- A. phô· mi· na·
father have field
"Father has a field."
- B. na· mi· klāa
field have salt
"The field has salt."
- C. phô· mi· na·klāa
father have salt-marsh
"Father has a salt marsh."

This is essentially the correct solution and certainly goes deeper than comments on head nouns and modifiers. However, there are several serious flaws in Warotamasikkkhadit's analysis. For one thing, his rules do not generate all the possible compounds in Thai. In particular, he has no rules at all which generate compounds whose first members are adverbials. This type represents a fairly large block of compounds, including ones like ro·nkhā·sāt 'slaughterhouse' to which we have already had occasion to refer.

For the compound types which Warotamasikkkhadit does cover, he does not attempt to maximize the generalities which can be observed in compound formation. The most crucial generality which he misses is the relationship between compound nouns and relative clauses. Warotamasikkkhadit does include a rule for the formation of relative clauses via a generalized transformation of the same type by which compounds are formed. But the compound rules start with two completely independent sentences, ignoring the mechanism already present in the grammar for embedding one sentence in the noun phrase of another.

Even within the approach which he takes, there are still generalities in compound types which Warotamasikkkhadit fails to exploit. One clear example involves three separate rules, all of which form compounds using sentences of the form Subject-Verb-Object, which are scattered among other compound-forming rules. One rule forms na.klāa 'salt marsh' in the manner illustrated above. The specification for the structure of which na.mi.klāa "The field has salt" is an example is:²⁷

$$N^1 + V_{m1} + N^2$$

V_{m1} is a special class of verbs of which mi 'to have' is the only member. As ad hoc as this may sound, there are good reasons for separating mi from other verbs in a grammar of Thai, so this is not the problem. But there is another rule which forms the compound khonna.n 'worker' (khon 'person', na.n 'work') from the sentence:

khon tham na.n
person do work

"The person does work."

The structural index for the embedded sentence in this rule is:²⁸

$$N^1 + V_t + N^2$$

It seems clear that the rule with V_{m1} is simply a special case of the rule with V_t . On the other hand, the model with which Warotamasikkkhadit was working had a serious problem of cross-classification of the type which led to the introduction of the notion of syntactic features into the theory. Thus, in this model, a verb could not be both a V_t and a V_{m1} , although

mi• is plainly a transitive verb. However, some of the generalization could easily have been captured by the use of a disjunction. The structural index for the embedded sentence of a rule which would collapse the above two rules would be:

$$N^1 + \left\{ \begin{matrix} V_t \\ V_{ml} \end{matrix} \right\} + N^2$$

The third rule of this general shape again includes the verb mi• 'to have'.²⁹ This rule makes the compound rótsă•mló• 'tricycle, pedicab' (rót 'vehicle', să•m 'three', ló• 'wheel') from the sentence rót mi• ló• să•m ló• "The vehicle has three wheels." The compounding rule is restricted to those nouns, like ló•, which serve as their own classifiers and deletes the pre-numeral occurrence of the nouns involved. But the deletion of the noun in situations like this is much more general than just in compound formation cases. If the rule which deletes the first occurrence of nouns in such cases is separated from the compound formation rules, as it seems it must anyway, the three rules could easily be collapsed, in Warotamasikkhadit's formulation, to:

$$N^1 + \left[\begin{matrix} V_t \\ V_{ml} \end{matrix} \right] + [(Nu)] + N^2$$

Since only nouns which serve as their own classifiers will have been deleted by the time this rule applies, only such nouns will ever meet the last half of the above rule and Warotamasikkhadit's restriction is preserved. This collapse of the three rules reveals the generality in the three cases, viz. they are all Subject-Verb-Object compounds.

3.0 Methodology and Data Sources. In the remainder of the chapter we shall analyze Thai noun compounds by type, making revisions of the compound forming rule and the rules which must follow it as we go along. The analysis is based on a large number of examples from various sources. The major source is Mary Haas' Thai-English Student's Dictionary.³⁰ I went through this dictionary, which is unusually rich in compound examples, taking a number of examples at random from each main entry which had more than three compounds as sub-entries, being careful to select any examples which appeared potentially troublesome. The definitions of almost all the examples cited in this chapter are from this source. To this was added the examples culled by Miss Udomphol from the government Thai-Thai Dictionary which were not already in the list.³¹ As a further check, I looked through a rather thorough Thai-Russian dictionary for examples which my analysis would not handle.³² Questionable examples were copied down and checked with my informant. In addition, there are a few examples which I collected in text from informants which do not appear in the dictionaries. The analysis below covers the vast majority of these examples. There is a scattering of compounds whose analysis resists the best efforts of me and my informants and probably should be considered single units.

4.0 Analysis of Compounds by Type.

4.1 Coordinate Compounds. Many Thai noun compounds are coordinate, that is, the meaning of the compound is a combination

of the meanings of the members. These compounds are derived from noun phrases which consist of coordinate nouns. Their origin is a general universal rule which forms coordinate structures from coordinate sentences. In the case of noun phrases, this rule would select the appropriate nouns from the two conjoined sentences and place them under the same noun phrase node with lé? 'and' between them. T lé? Del can delete this word and leave the two nouns. Hence, our first approximation to the general noun compounding rule is simply:

$$\begin{array}{c} \text{GNC } \langle \text{N } \text{N} \rangle_{\text{NP}} \\ \quad \quad \quad \begin{array}{c} 1 \quad 2 \\ 1 > 2 \end{array} \end{array}$$

This takes the second noun of a noun phrase which consists of only two nouns and inserts it under the node dominating the first. The examples of coordinate compounds in Thai are the following:³³

pètkày 'barnyard fowl'

pèt 'duck'

kày 'chicken'

phī·nó·n 'brothers and sisters'

phī 'older sibling'

nó·n 'younger sibling'

phī·pa·ná?a· 'relatives'³⁴

phī·pa· 'older sibling and father's older sister'

ná?a· 'mother's younger sister and mother's older sister'

kamlanre·n 'power'

kamlan 'strength'

re·n 'power'

dinfá·ʔa·ka·t 'climate'

dinfá· 'earth and sky'

ʔa·ka·t 'air'

phũamia 'married couple'

phũa 'husband'

mia 'wife'

lomfá·ʔa·kà·t 'weather, climate'

lom 'wind'

fá· 'sky'

ʔa·kà·t 'air'

lamklôn 'gun barrel'

lam 'classifier for cylindrical objects'

klôn 'pipe'

siannă·m 'enemies of society'

sian 'splinter'

nă·m 'thorn'

ple·wfay 'flame'

ple·w 'flame'

fay 'fire'

4.2 Compounds Derived from Noun Phrases with Relative
Clauses. Coordinate compounds are unique in several ways. Not the least of these is the fact that they are the only type not derived via relative clauses. Of those which are derived from relative clauses, there are two major types. In all compounds, there is one sentence embedded in the noun phrase involved which reveals the grammatical relationships between the members of the compound. In one of the major subdivisions, this sentence is embedded directly in a relative clause which modifies the head noun of the compound. An example of this is sàtpà· 'wild animal' (sàt 'animal', pà· 'forest'). The sentence

reflecting the relationship between sàt and pà· is:

sàt yù· nay pà·
animal be-located in forest

"The animal lives in a forest."

The derivation of this compound, up to compound formation itself, would be:³⁵

BASE: <sàt <sàt ?a·săyyù· nay pà· >S>NP
 animal animal dwell in forest

FIRST CYCLE: No operations.

SECOND CYCLE:

T Rel Inst:

<sàt <thi· sàt ?a·săyyù· nay pà· >S>NP
 animal which animal dwell in forest

T Red N Del:

<sàt <thi· ?a·săyyù· nay pà· >S>NP
 animal which dwell in forest

T Rel Pro Pre: Applies vacuously.

T Pron Del:

<sàt <?a·săyyù· nay pà· >VP>NP
 animal dwell in forest

The other major subdivision involves the diagnostic sentence being embedded at a deeper layer, specifically as the nominalized object of the preposition sămràp 'for' in the verb phrase cháy sămràp 'to be used for'. An example of this type of compound is să·yka·nbin 'airline' (să·y 'line', ka·nbin 'flying'). Its derivation is considerably more complicated and is given below, up to compound formation:

BASE:

<să.y <khon cháy sămrap <ka.n <ka.n pen <<khon bin
 line person use for activity activity be person fly
dây să.y >NP>S>NP să.y >NP ³⁶
 with line line line

FIRST CYCLE: No operations.

SECOND CYCLE: No operations.

THIRD CYCLE:

T Am Vb:

<să.y <să.y cháy sămrap <ka.n <ka.n pen <khon bin
 line line be-used for activity activity be person fly
dây să.y >NP>S>NP>S>NP
 with line line

T Rel Inst:

<să.y <să.y cháy sămrap <ka.n <thi. ka.n pen
 line line be-used for activity which activity be
<khon bin dây să.y >NP>S>NP>S>NP
 person fly with line

T Red N Del:

<să.y <să.y cháy sămrap <ka.n <thi. pen <khon bin
 line line be-used for activity which be person fly
dây >NP>S>NP>S>NP
 with

T Rel Pro Pre: Applies vacuously.

T Pron Del:

<să.y <să.y cháy sămrap <ka.n <pen <khon bin
 line line be-used for activity be person fly
dây >NP>VP>NP>S>NP
 with

T App:

<să.y <să.y cháy sămrap <ka.n <khon bin dây >NP>NP>S>NP
 line line be-used for activity person fly with

T Nom:

<să·y <să·y cháy sămrap <ka·nbin dây khố·n khon >NP>S>NP
line line be-used for flying with of person

T Pro Del:

<să·y <să·y cháy sămrap <ka·nbin dây khố·n >NP>S>NP
line line be-used for flying with of

T Is Prep Del:

<să·y <să·y cháy sămrap <ka·nbin >N>S>NP
line line be-used for flying

FOURTH CYCLE:

T Rel Inst:

<să·y <thi· să·y cháy sămrap <ka·nbin >N>S>NP
line which line be-used for flying

T Red Nom Del:

<să·y <thi· cháy sămrap <ka·nbin >N>S>NP
line which be-used for flying

T cháy Del:

<să·y thi· sămrap <ka·nbin >N>NP
line which for flying

T Pron Del:

<să·y sămrap <ka·nbin >N>NP
line for flying

T sămrap Del:

<să·y <ka·nbin >N>NP
line flying

It is apparent that both types of derivation are needed.

să·yka·nbin could not be derived directly from a relative clause since there is no *să·y thi· ka·nbin 'line which flying'. Similarly, while it would not be strictly ungrammatical, săt thi· sămrap ka·n?a·săyyù· nay pà· 'an animal which is for living in

the forest' is clearly not the correct source for sàtpà. While the cháy sāmrap derivation seems complicated, it must be remembered that all the rules involved are needed in a grammar of Thai anyway. Denying this derivation for compounds, even if it were possible, would not simplify the grammar at all; it would merely prevent the rules from applying in a certain way.

On the other hand, most compounds are not as clear-cut as to which derivation is appropriate. In many cases, my informant would accept either derivation. Accordingly, in our lists of examples, we will give an example which can be derived via the cháy sāmrap derivation and one which can be derived directly from a relative clause. We will not specify derivation for the rest but place them in a single list with the understanding that many are ambiguous in this way.

4.2.1 Compounds with Subject as Head.

4.2.1.1 Predicate Nominative as Second Member. One type of compound for which the cháy sāmrap derivation will never be appropriate are those whose second member is a predicate nominative. A common use of this derivation is to generate compounds whose first member is a noun of generic meaning and whose second member has a more specific meaning. Another type of compound so generated are kinship terms. An interesting use of this mechanism is the generation of compounds whose first member has a general meaning and whose second member is a loan-word from another language, usually English. An example is

rótté·ksĩ 'taxi' (rót 'vehicle', té·ksĩ 'taxi'). A sample derivation demonstrates the origin of hônkhrua 'kitchen'.

BASE: <hôn <hôn pen khrua >_S>_{NP}
 room room be kitchen

FIRST CYCLE: No operations.

SECOND CYCLE:

T Rel Inst:

<hôn <thĩ· hôn pen khrua >_S>_{NP}
 room which room be kitchen

T Red N Del:

<hôn <thĩ· pen khrua >_S>_{NP}
 room which be kitchen

T Rel Pro Pre: Applies vacuously

T Pron Del:

<hôn <pen khrua >_{VP}>_{NP}
 room be kitchen

T App:

<hôn khrua >_{NP}
 room kitchen

Notice that the derivation yields a structure exactly like that which is obtained in the case of coordinate compounds. As a result, GNC needs no revision and in its present form yields hônkhrua 'kitchen'.

Examples of this type of compound are the following:

hônkhrua 'kitchen'

hôn 'room'

khrua 'kitchen'

phĩ·cha·y 'elder brother'

phĩ 'older sibling'

cha·y 'human male'

kò?lan̩ka 'Island of Ceylon'

kò? 'island'

lan̩ka 'Ceylon'

khî·khâ 'slave'

khî 'excrement'

khâ 'servant'

na·nwí·cay 'research work'

na·n 'work'

(ka·n)wí·cay 'research'³⁷

khanchat 'large fan'

khan 'classifier for long-handled objects'

chat 'fan'

ca·nphí 'flying saucer'

ca·n 'plate, dish'

phí 'ghost, spirit'

to·nkhâm 'evening'

to·m 'classifier for sections of time or space'

khâm 'evening'

nákkawí 'poet'

nák 'expert'³⁸

kawí 'poet'

phle·nsawín 'swing music'

phle·n 'song'

sawín 'swing' (from English)

4.2.1.2 Verb and Object as Second and Third Members.

Unlike English, Thai has a number of compounds which have three members at the same level of embedding. Many of these are Subject-Verb-Object compounds derived ultimately from sentences consisting of the same sequence of words. The derivation of khri·an̩khum?a·kà·t 'air conditioner' will illustrate

the cháy săm ràp derivation type.

BASE:

<khri̯an <khon cháy săm ràp <ka·n <ka·n pen <<khri̯an
device person use for activity activity be device

khum ?a·kà·t >S>NP>S>NP khri̯an >S>NP
watch-over air device

FIRST CYCLE: No operations.

SECOND CYCLE: No operations.

THIRD CYCLE:

T Am Vb:

<khri̯an <khri̯an cháy săm ràp <ka·n <ka·n pen <khri̯an
device device be-used for activity activity be device

khum ?a·kà·t >NP>S>NP>S>NP
watch-over air

T Rel Inst:

<khri̯an <khri̯an cháy săm ràp <ka·n <thi· ka·n pen
device device be-used for activity which activity be

<khri̯an khum ?a·kà·t >NP>S>NP>S>NP
device watch-over air

T Red N Del:

<khri̯an <khri̯an cháy săm ràp <ka·n <thi· pen <khum
device device be-used for activity which be watch-over

?a·kà·t >NP>S>NP>S>NP
air

T Rel Pro Pre: Applies vacuously.

T Pron Del:

<khri̯an <khri̯an cháy săm ràp <ka·n <pen <khum
device device be-used for activity be watch-over

?a·kà·t >NP>VP>NP>S>NP
air

T App:

~~<khriān~~ ~~<khriān~~ cháy sămrap ~~<ka·n~~ ~~<khum~~
 device device be-used for activity watch-over
?a·kà·t
 air NP>NP>S>NP

T Nom:

~~<khriān~~ ~~<khriān~~ cháy sămrap ~~<ka·nkhum~~ ?a·kà·t
 device device be-used for watching-over air N>S>NP

FOURTH CYCLE:

T Rel Inst:

~~<khriān~~ ~~<thi·~~ khriān cháy sămrap ~~<ka·nkhum~~ ?a·kà·t
 device which device be-used for watching-over air N>S>NP

T Red N Del:

~~<khriān~~ ~~<thi·~~ cháy sămrap ~~<ka·nkhum~~ ?a·kà·t
 device which be-used for watching-over air N>S>NP

T cháy Del:

~~<khriān~~ thi· sămrap ~~<ka·nkhum~~ ?a·kà·t
 device which for watching-over air N>NP

T Pron Del:

~~<khriān~~ sămrap ~~<ka·nkhum~~ ?a·kà·t
 device for watching-over air N>NP

T sămrap Del:

~~<khriān~~ ~~<ka·nkhum~~ ?a·kà·t
 device watching-over air N>NP

T ka·n Del:

~~<khriān~~ ~~<khum~~ ?a·kà·t
 device watching-over air N>NP

The section of this process which differs from the derivation of să·yka·nbin 'airline' is that part which follows the deletion of the subject of the innermost embedded sentence by T Red

N Del in the third cycle. The results of this step affect the application of T Nom and make unnecessary the application of T Pro Del and T Is Prep Del. The striking fact about the structure in the last line above is that it meets the structural requirements for GNC as it now stands, in spite of the apparent wide difference between Subject-Verb-Object compounds and Subject-Predicate Nominative and coordinate compounds. The head noun khriān is obviously a noun and so, by virtue of T Nom, is khum-?a·kà·t. Thus the structure above is a case of $\langle N N \rangle_{NP}$.

The compound khè·tplò·tthahă·n 'demilitarized zone' illustrates the more direct method of derivation.

BASE:

$\langle \frac{\text{khè·t}}{\text{area}} \frac{\text{khè·t}}{\text{area}} \frac{\text{plò·t}}{\text{lack}} \frac{\text{thahă·n}}{\text{soldier}} \rangle_S \rangle_{NP}$

FIRST CYCLE: No operations.

SECOND CYCLE:

T Rel Inst:

$\langle \frac{\text{khè·t}}{\text{area}} \frac{\text{thi·}}{\text{which}} \frac{\text{khè·t}}{\text{area}} \frac{\text{plò·t}}{\text{lack}} \frac{\text{thahă·n}}{\text{soldier}} \rangle_S \rangle_{NP}$

T Red N Del:

$\langle \frac{\text{khè·t}}{\text{area}} \frac{\text{thi·}}{\text{which}} \frac{\text{plò·t}}{\text{lack}} \frac{\text{thahă·n}}{\text{soldier}} \rangle_S \rangle_{NP}$

T Pron Del:

$\langle \frac{\text{khè·t}}{\text{area}} \frac{\text{plò·t}}{\text{lack}} \frac{\text{thahă·n}}{\text{soldier}} \rangle_{VP} \rangle_{NP}$

The structure generated by this method does not meet the requirements for the application of GNC and will force a modification of that rule. Thus it can be seen that the simplicity

of this derivation is more apparent than real, since it forces a complication in another part of the grammar. The new formulation of GNC is:

$$\begin{array}{c} \text{GNC} \quad \langle \underset{1}{\text{N}} \quad (\text{V}) \quad \underset{2}{\text{N}} \rangle_{\text{NP}} \\ \rightarrow 1>2 \end{array}$$

Examples of this type are:

khonkhă·ytũa 'ticket agent'

khon 'person'

khă·y 'to sell'

tũa 'ticket'

khê·tplồ·tthahă·n 'demilitarized zone'

khê·t 'area'

plồ·t 'to lack'

thahă·n 'soldier'

bàtpracamtua 'identification card'

bàt 'card'

pracam 'to represent'

tua 'self'

phâ·kanpian 'apron'

phâ· 'cloth'

kan 'to prevent'

(ka·n)pian 'getting dirty'

phê·tphadunkhan 'obstetrician'

phê·t 'physician'

phadun 'to support'

khan 'pregnancy'

mỗ·sổ·nsà·sană· 'missionary'

mỗ· 'doctor'

sổ·n 'to teach'

gà·sană· 'religion'

dèklíankè? 'shepherd boy'

dèk 'child'

lían 'to care for'

kè? 'sheep'

khonlíandèk 'foster parent'

khon 'person'

lían 'to care for'

dèk 'child'

châ·ntàtphôm 'barber'

châ·n 'artisan'

tàt 'to cut'

phôm 'hair'

dèk?ommí· 'infant'

dèk 'child'

?om 'to keep in the mouth'

mí· 'hand'

4.2.1.3 Verb as Second Member. More common than Sub-

ject-Verb-Object compounds are compounds composed of the subject and verb alone. These are of three subtypes: a) compounds in which the underlying object has been deleted, b) compounds in which the verb is intransitive, but non-adjectival, and c) compounds in which the verb is adjectival. Only the first two subtypes are amenable to cháy sǎmráp derivation.

The type a) compounds are most like the Subject-Verb-Object type discussed above. The word meaning 'alarm clock' na·líka·plùk, illustrates the cháy sǎmráp derivation.

BASE:

<na·líka· <khon cháy sǎmráp <ka·n <ka·n pen <<na·líka·
clock person use for activity activity be clock

plùk khon >S>NP>S>NP na·líka·>S>NP
arouse person clock

GNC applies to this structure to form na.lika.plùkkhon, a non-occurring compound. In order to generate na.lika.plùk, we will have to add a rule which applies to structures generated by GNC to delete unwanted members in cases like this. The rule is of this form:

This optional rule deletes the third member of a noun compound of the form Noun-Verb-Noun if the third member is a Pro-noun like khon 'person'.

BASEs

The derivation is parallel to that of khè·tplò·tthahă·n 'demilitarized zone' and ends with:

GNC readily converts this into khonɬə̌·khǝ̌·n. Since khǝ̌·n 'thing' has the feature [+Pro], it is deleted by NCD 1, giving khonɬə̌·.

Examples of subtype a) of Subject-Verb compounds are:

khonsí 'customer'

khon 'person'

sí 'to buy'

nákkhian 'writer'

nák 'expert'

khian 'to write'

nákcí 'hold-up man'

nák 'expert'

cí 'to point, hold-up'

na·mfě·n 'pseudonym'

na·m 'name'

fě·n 'to hide, conceal'

na·lika·plùk 'alarm clock'

na·lika 'clock'

plùk 'to arouse'

phû·kháp 'operator'

phû 'person'

kháp 'to drive'

phô·lian 'stepfather'

phô 'father'

lian 'to care for'

mố·du 'fortune-teller'

mố 'doctor'

du 'to see'

phû·títta·m 'an aide who travels with his superior'

phû 'person'

títta·m 'to accompany'

fă·khrô·p 'dish cover'

fă 'lid'

khrô·p 'to cover up'

Subtype b) of Subject-Verb compounds is similar to subtype a) except that the verb is intransitive. Both cháy sảmráp and direct relative clause derivations are necessary. The cháy sảmráp derivation can be illustrated with the derivation of má·khèn 'race horse'.

BASE:

má· <khon <cháy sảmráp <ka·n <ka·n pen <<má·
horse person use for activity activity be horse
khèn >S>NP>S>NP má· >S>NP
compete horse

The derivation of má·khèn is parallel to that of na·lika·plùk 'alarm clock' except that there is no underlying object. The process ends with:

<má· <khèn >N>NP
horse compete

GNC converts this directly into má·khèn 'race horse'.

The direct derivation for subtype b) compounds can be illustrated with ríabin 'airplane'.

BASE: <ríá <ríá bin>S>NP
boat boat fly

The derivation parallels that of khonsí· 'customer' except for the lack of an object. The last line contains the structure:

<ríá <bin>VP>NP
boat fly

GNC in its present form will not operate on this structure. It seems to have developed that GNC operates on two kinds of structures. 1) <N N>NP and <N V (N)>NP. Accordingly, we shall

introduce the following disjunction into GNC:

$$\begin{array}{c} \text{GNC} < \text{N} \left\{ \begin{array}{c} \text{N} \\ \text{V} \end{array} \right. \left. \begin{array}{c} \text{(N)} \\ \text{(N)} \end{array} \right\} > \text{NP} \\ \rightarrow \begin{array}{cc} 1 & 2 \\ 1 > 2 \end{array} \end{array}$$

With this modification, GNC will generate riabin 'airplane' from the above structure. Examples of subtype b) compounds are those below:

má·khèn 'race horse'

má· 'horse'

khèn 'to compete'

riabin 'airplane'

ria 'boat'

bin 'to fly'

khampen 'live syllable'

kham 'word'

pen 'to be, exist, live'

ca·nbin 'flying saucer'

ca·n 'plate, dish'

bin 'to fly'

cha·wpramon 'fisherman'

cha·w 'dweller in'³⁹

pramon 'to fish'

dinrabè·t 'explosive powder'

din 'earth'

rabè·t 'to explode'

ne·nta·y 'rubber check'

ne·n 'money'

ta·y 'to die'

ná·mkha·n 'dew'

ná·n 'water'

kha·n 'to remain'

nákwíca·n 'commentator'
nák 'expert',⁴⁰
wíca·n 'to comment'

hínŋô·k 'stalagmite'
hín 'rock'
ŋô·k 'to sprout, bud, grow'

Compounds of subtype c) differ from subtype b) only in that the verb in the innermost embedded sentence is an adjectival verb and that cháy sāmrap derivation is impossible. The derivation of words like kracòkwáw 'concave lens' (kracòk 'lense', wáw 'to be concave') is exactly like that of riabin 'airplane'. Examples of this kind are:

kracòkwáw 'concave lens'
kracòk 'lense'
wáw 'to be concave'

khôn?ò·n 'down'
khôn 'hair, feathers'
?ò·n 'to be young, tender'

khamyô· 'abbreviation'
kham 'word'
yô· 'to be short'

ná·nhô·m 'perfume'
ná·m 'water, fluid'
hō·m 'to be fragrant'

ná·nǎ·t 'dark horse (in politics)' (probably a loan translation from English)
nǎ· 'horse'
nǎ·t 'to be dark'

nianǎ·y 'minor or additional wife'
nǎ 'wife'
nǎ·y 'to be new'

ya·de·n 'mercurochrome'

ya· 'medicine'

de·n 'to be red'

lũ·kʔò·n 'infant, young animal'

lũ·k 'offspring, child'

ʔò·n 'to be young, tender'

mổ·thàn 'quack'

mổ· 'doctor'

thàn 'to be wild, unauthorized'

phũ·ʔá·n 'others'

phũ· 'person'

ʔá·n 'to be other'

4.2.1.4 Object as the Second Member. Just as subtype a) of the Subject-Verb compounds differed from the Subject-Verb-Object type by deletion of the underlying object, so there is a set of compounds of the shape Subject-Object, in which the verb has similarly been deleted. The cháy sảmráp as well as the relative clause method of derivation is again relevant. The cháy sảmráp derivation can be illustrated with the compound wuanom 'milch cow'.

BASE:

<wua <khon cháy sảmráp <ka·n <ka·n pen <<wua hây
cow person use for activity activity be cow give

nom >S>NP>S> wua>S>NP
milk cow

The derivation from this point is the same as that of khrián-khumʔa·kà·t 'air conditioner'. The last line before application of GNC is:

$\langle \text{wua} \langle \text{hây} \text{ nom} \rangle_N \rangle_{NP}$
 cow give milk

GNC converts this structure into wuahâynom. It is now necessary to add to our grammar a rule which optionally deletes verbs from noun compounds of this kind. NCD 2 is designed to do this.

NCD 2 $\langle N \ V \ N \rangle_N$
 $\begin{matrix} 1 & 2 & 3 \\ \rightarrow & 1 & \emptyset & 3 \end{matrix}$

The direct derivation from relative clauses can be illustrated using rô·kcît 'mental illness'.

BASE:

$\langle \text{rô·k} \langle \text{rô·k} \text{ pùay} \text{ tha·n} \text{ cît} \rangle_S \rangle_{NP}$
 disease disease affect along mind

A derivation process parallel to that of khètplò·tthahă·n 'demilitarized zone' except for the presence of the preposition tha·n 'along', ends with the structure:

$\langle \text{rô·k} \langle \text{pùay} \text{ tha·n} \text{ cît} \rangle_{VP} \rangle_{NP}$
 disease affect along mind

In order for the General Noun Compounding rule to operate on this structure, it will have to be revised again to include cases in which a verb is of the type that has a preposition inserted between it and its object. The new form of GNC is:

GNC $\langle N \left\{ \begin{matrix} N \\ V(\text{Prp})(N) \end{matrix} \right\} \rangle_{NP}$
 $\begin{matrix} 1 & 2 \\ \rightarrow & 1 > 2 \end{matrix}$

In this form, GNC will generate rô·kpùaytha·ncît. In order to get rô·kcît, NCD 2 will have to be revised to delete prepositions

as well. In addition, we notice that there are no four-member compounds of the form Noun+Verb+Preposition+Noun, so we must make NCD 2 obligatory if a preposition is present. The new form of NCD 2 is :

NCD 2	<N	V	(Prp)	N>	N	Obligatory if 3 is present.
	1	2	3	4		
→	1	∅	∅	4		

It may seem that the second half of the disjunction in GNC should refer to Verb Phrase instead of the three constituents above. However, there are many Verb Phrase structures which do not appear in compound derivations, so it is more efficient to refer to those that do, specifically.

rô·kcît is one of many diseases which are named with compounds with rô·k 'disease' and khây 'fever, sickness'. The structure of some of these is not as transparent as that of most compounds at first glance. Some of these troublesome compounds are those below:

khây·râ·ksà·t 'typhoid'

khây 'fever, sickness'

râ·ksà·t 'to vomit'

rô·k·ta·de·n 'conjunctivitis'

rô·k 'disease'

ta· 'eye'

de·n 'to be red'

The cases in which the compound has three members of which the second and third are Noun+Verb in apparent Subject-Verb relationship can be handled by the grammar as it exists. If we allow this part to be compounded first in the same manner

as kracòkwáw 'concave lense', then it has the structure <ta·de·n>_N 'red-eye'. If this structure is available, rô·kta·de·n can be derived by existing rules via a noun phrase of the form:

rô·k thi· tham ta·de·n
disease which cause red-eye

rô·kta·de·n will be generated by the application of GNC as outlined just above followed by the operation of NCD 2.

khây·râ·ksà·t 'typhoid' can be handled as easily by a slightly different route. The base for this compound will be:

<khây <khây tham <ka·n <ka·n pen <<khon
sickness sickness cause activity activity be person
râ·ksà·t >S>NP>S>NP>S>NP
vomit

Rules already familiar to us will generate:

khây thi· tham ka·nrâ·ksà·t
sickness which cause vomiting

Application of T Pron Del, T ka·n Del, GNC and NCD 2 in the ordinary way will then generate khây·râ·ksà·t. In neither case is it necessary to add rules to our grammar.

A comparison of NCD 1 and NCD 2 might lead one to expect that they might be made more analogous if NCD 2 referred to a constituent like $\begin{bmatrix} +V \\ +Pro \end{bmatrix}$. This would mean that not just any verb, but a few certain verbs of general meaning only could be deleted in compounding. There is considerable evidence, however, that this is not possible. While there are a number of compounds which can be generated on the basis of a few

general verbs, there are many others which demand more specific verbs. There are a number of compounds which seem to be generable from sentences of which the main verb is mi 'to have', as we discussed earlier. For example thùafàkya·w 'long-podded cowpea' (thùà 'pea', fàkya·w 'long pod') seems to be derivable from thùà thi· mi· fàkya·w 'a pea which has a long pod'. Others in large numbers can be derived from kław káp 'to deal with'. One such compound is thabianrót 'car registration' (thabian 'registration', rót 'car'). Another common verb which might be a candidate for the feature [+Pro] is mian káp 'to resemble'. The word for 'gold leaf' tho·nkhample·w (tho·n·kham 'gold', ple·w 'flame') can be derived from tho·nkham thi· mian káp ple·w 'gold which resembles flames'.

To begin with, there seems to be little reason to group these three words together with a single feature except for the fact that they all figure in the derivation of a number of compounds. But in fact, these three, even if joined by a few more such verbs, would be insufficient to account for all Subject-Object compounds in Thai. To generate compounds with rô·k, we have seen already that we needed tham 'to cause' and pùay tha·n 'to affect'. To generate wuanom, we needed hây 'to give'.⁴¹ The list of verbs grows as we look at more and more compounds. It seems that we must allow verbs to be freely deleted in compounds by NCD 2, and accept the result that compounds are potentially ambiguous in indefinitely many ways, if they involve a deleted verb.⁴²

We now consider a group of compounds which presents quite a puzzle. Among these, we find sănă.myâ 'lawn' (sănă.m 'field', yâ 'grass'), mũ.să.mchán 'bacon' (mũ 'pork', să.m 'three', chán 'layer'), and khâ.wplăk 'unhusked rice' (khâ.w 'rice', plăk 'husk'). It appears that these are a special class of Subject-Object compounds in which the underlying verb is mi 'to have'. Thus these compounds would have the following ultimate sources:

sănă.myâ from sănă.m mi yâ
grass field field have grass

mũ.să.mchán from mũ mi să.m chán
three layer pork pork have three layer

khâ.wplăk from khâ.w mi plăk
husk rice rice have husk

Indeed, for mũ.să.mchán, Mr. Sirisamphan gave me:

mũ suan thi mi să.m chán
pork part which have three layer

as a definition for mũ.să.mchán, but for khâ.wplăk, he gave me the startling source sentence:

khâ.w thi pen thăn plăk
rice which be all husk

Furthermore, many of these compounds have structures which are unique in that they are composed of Noun-Number-Classifier, as mũ.să.mchán and sàtsi.thăw 'four-legged animal'.

An examination of some other strange phenomena in Thai seemed to have a bearing on the solution. Copula sentences in Thai have a number of peculiarities from the point of view of an English speaker. First, there is an ordinary, expected

kind of sentence:

khô·pkhrua phôm pen khô·pkhrua lèk
family me be family small

"My family is a small one."

Then there are a few of the type:

khon pen thay
person be Thai

"The person is Thai."

This is a little odd, because descriptive adjectives are stative verbs in Thai, and as a noun they rarely occurs except as a member of a compound like khonthay 'Thai, Thai person', pràthê·tthey 'Thailand', and mə·wthay 'Siamese cat'. Then, there is another interesting set:

khăw mák cà? khian bò·k wáy wā· khon pen phê·t ʔaray
they likely will write tell keep that person is sex what

"They will probably write down what sex the person is."

This type of sentence has only marginal equivalents in English. We may say "What sex is this puppy?", but I suspect many English speakers would be uneasy about it. The Thai sentences seem quite natural. Another group is even more odd:

nay lò·k ní· mí· nə·y khon thí· mây khe·y pen khây
in world this have few person who not ever be sickness

rí· pen rô·k le·y
or be disease at-all

"In this world there are few people who have never been sick or ill."

khon thí· pen sò·t yan mây mí· phanráya.⁴⁴
person who be single-state still not have wife

"A person who is single doesn't have a wife yet."

pen fan
be tooth

"It's serrated."

In these sentences, the predicate noun seems to be completely incompatible with the subject to a degree which is entirely unmatched in English. The subject is somehow understood to partake of the quality of the predicate nominal. The predicate nominal need not be a single noun, as in:

pen ko·n số·n ko·n
be heap two Clf.

"It's a whole lot."

Having discovered that Thai has stative verbs, it is a bit disturbing to find sentences like:

ke· mi· phủ·chả·y ?i·k khon ma· yê·n fe·n ke·
he have man other Clf. come snatch girlfriend him

pay lê·y pen bâ·
go hence be be-crazy

"Another man came and took his girlfriend away; that's why he's crazy."

pen bây
be gesture-silently

"He's mute."

pen pay dây
be go can

"It's possible."

A look at one more example will provide the key to all these puzzling cases:

pen thi· pho·cay
be what satisfy

"It's satisfactory."

It is obvious that in the cases in which the verb follows pen, there was a relative pronoun which has been deleted by T Pron Del. Hence pen bâ· 'be be-crazy' was pen thi· bân 'be (someone)

who is crazy', pen bây 'be gesture-silently' was pen thi• bây
'be (someone) who gestures silently' (because he is mute), etc.

It still remains to specify just how such sentences develop. I would like to suggest that the deep structure string for such sentences has the form:

N pen N thi• N Pred
1 2 3 4 5 6

where 3=5 (so that 5 is deleted by T Red N Del), and that 3 is a classifier compatible with 1 or the same as 1. It will be remembered that all classifiers are also nouns. Under these conditions, 3 may be deleted. Rules already in the grammar will effect the other necessary deletions. The new rule may be stated as:

A. $\langle\langle X Y Z \rangle_{NP} \text{ pen } N \text{ thi} \cdot VP \rangle_S$ Where 2 is $\begin{bmatrix} +N \\ +(5) \text{ Nu} \end{bmatrix}$
 1 2 3 4 5 6 7 or 2=5.
 → 1 2 3 4 Ø 6 7

Thus the origin of pen bâ• is:

*khâw pen khon thi• khon bâ•
 he be person which person be-crazy

The cases which involve predicate nouns incompatible with the subject are still not solved, however. But it now seems likely that there is a special mi•-deletion rule which operates in these copula sentences. So, we propose that a sentence like khon pen rô•k 'The person is sick' be given the structure associated with:

khon pen khon thi• khon mi• rô•k
 person be person which person have disease

The rules we now have will give us:

khon pen mi• rô•k
 person be have disease

A transformation of the form of B will give us the desired sentences:

B. NP pen mi• NP
 1 2 3 4
 → 1 2 ∅ 4

We are left with a situation in which mi• 'to have' is involved in sentences which are odd in two ways. First they convert to sentences with pen 'to be' and secondly, they provide the basis for compounds with numbers. It would seem desirable to account for both these phenomena at once. The way to do this would seem to be to allow sentences with mi• to be converted to sentences with pen and to allow the compounds to be derived from the result as Subject-Predicate Nominal compounds, not as Subject-Object compounds. This would indeed be the answer if it were the case that Subject-Predicate Nominal compounds, but not Subject-Object compounds, could have numbers in their second members. However, this is not the case, even for those compounds which led us to explore the possibility to begin with. In testing sets of phrases with Mr. Sirisamphan, the following sets were completely acceptable:

ná•m thi• mi• khlo•n
 water which have mud

ná•m thi• pen khlo•n
 water which be mud

khon thi• mi• khây
 person who have sickness

khon thi• pen khây
 person who be sickness

khâ•w thi• mi• plâk
 rice which have husk

khâ•w thi• pen plâk
 rice which be husk

Rules which we now have will convert either alternative into

ná·m̀khlo·n 'muddy water', khonkhây 'patient', and khâ·wpḷak 'unhusked rice', respectively. But in the crucial cases, the second alternative proved unacceptable.

rôtcàkraya·n thî· mî· *rôtcàkraya·n thî· pen
bicycle which have bicycle which be

să·m lô·
three wheel

să·m lô·
three wheel

săt thî· mî· số·n thá·w
animal which have two foot

*săt thî· pen số·n thá·w⁴⁵
animal which be two foot

As a result, only the first of the two possibilities are available as sources for rôtcàkraya·nsă·mlô· 'pedicab' and săt·số·n·thá·w 'two-legged animal'. There is nothing to be gained, then by allowing the conversion of mî· sentences to pen sentences to be a part of the compound derivation process. There seems to be no more elegant way to account for the fact that numbers can appear in only those compounds which have mî· 'to have' in the underlying sentence. Hence, GNC must be revised to read:⁴⁶

$$\begin{array}{c} \text{GNC} <N & \left\{ \begin{array}{c} N \\ V \text{ (Prp) (N)} \\ \underline{mî·} \text{ (Nu) } N \end{array} \right\} >NP \\ & \begin{array}{cc} 1 & 2 \\ \longrightarrow & 1>2 \end{array} \end{array}$$

We are now ready to give examples of Subject-Object compounds. It will be interesting to list these in three lists; those deriving from sentences with mî· 'to have', those deriving from sentences with ṃan kàp 'to resemble', and those deriving from sentences with other verbs. Examples of Subject-Object compounds from sentences with mî· 'to have':

kâynuan 'turkey'

kây 'chicken'

nuan 'proboscis'

câ·wkhǎ·nbâ·n 'landlord'

câ·wkhǎ·n 'proprietor'

bâ·n 'house'

thũnyâ· 'prairie'

thũn 'field, meadow'

yâ· 'grass'

mékka·si·nphâ·p 'picture magazine'

mékka·si·n 'magazine' (from English)

phâ·p 'picture'

rótsǎ·mló· 'pedicab'

rót 'vehicle'

sǎ·m 'three'

ló· 'wheel'

sàtsǎ·nthá·w 'two-legged animal'

sát 'animal'

sǎ·nthá·w 'two feet'

ná·mkhlo·n 'muddy water'

ná·m 'water'

khlo·n 'mud'

khonkhây 'patient'

khon 'person'

khây 'sickness'

khâ·wplǎk 'unhusked rice'

khâ·w 'rice'

plǎk 'husk'

rótcàkraya·nǎ·mló· 'pedicab'

rótcàkraya·n 'bicycle'

sǎ·mló· 'three wheels'

Examples of Subject-Object compounds from sentences with mian
kàp 'to resemble':

khanõmpiakpu·n 'a kind of pudding'

khanõmpiak 'wet dainty'

pu·n 'mortar'

khăwkin 'antler'

khăw 'horn'

kin 'twig'

so·da·fay 'caustic soda'

so·da· 'soda' (from English)

fay 'fire'

tho·nkhample·w 'gold leaf'

tho·nkhām 'gold'

ple·w 'flame'

nókkê·w 'parrot'

nók 'bird'

kê·w 'crystal'

pla·khēm 'needle fish'

pla· 'fish'

khēm 'needle'

phríkkhī·nũ· 'bird pepper'

phrík 'pepper'

khī·nũ· 'mouse droppings'

kradà·tkê·w 'cellophane'

kradà·t 'paper'

kê·w 'glass'

wonlép 'parenthesis'

won 'circle'

lép 'fingernail'

mí·khwǎ·y 'a kind of large bear'

mí· 'bear'

khwǎ·y 'water buffalo'

Examples of Subject-Object compounds from sentences with other verbs:

krasuanka·ntà·nprathê·t 'Ministry of Foreign Affairs'⁴⁷

krasuan 'ministry'

ka·ntà·nprathê·t 'matters of other countries'

khây·râ·ksà·t 'typhoid'

khây 'sickness'

râ·ksà·t 'to vomit'

rô·kta·de·n 'conjunctivitis'

rô·k 'disease'

ta·de·n 'red-eye'

wuanom 'milch cow'

wua 'cow'

nom 'milk'

thabianrót 'car registration'

thabian 'registration'

rót 'vehicle'

na·yphon 'general'

na·y 'master'

phon 'troops'

rabê·tparama·nu· 'atomic bomb'

rabê·t 'bomb'

parama·nu· 'atom'

rô·kfi·nfará·y 'cancer'

rô·k 'disease'

fi·nfará·y 'viscious flesh boil'

wisawakamfayfá 'electrical engineering'

wisawakam 'engineering'

fayfá 'electricity'

châ·nfay 'fireman'

châ·n 'artisan'

fay 'fire'

4.2.1.5 Verb and Adverbial as Second and Third Members.

There are a few three-member compounds reflecting the subject, verb and adverbial constituents of the underlying sentence. Compared to the Subject-Verb-Object type, there are surprisingly few of these. However, they are easily generable by the grammar of compounds under discussion. The compound riade·n-samùt 'ocean-going steamer' is representative of a compound derived from a cháy sāmrap sentence.

BASE:

<riá <khon cháy sāmrap <ka·n <ka·n pen <<riá də·n
boat person use for activity activity be boat walk

nay samùt >S>NP>S>NP riá >S>NP
in ocean boat

FIRST CYCLE: No operations.

SECOND CYCLE: No operations.

THIRD CYCLE:

T Am Vb:

<riá <riá cháy sāmrap <ka·n <ka·n pen <riá də·n
boat boat be-used for activity activity be boat walk

nay samùt >NP>S>NP>S>NP
in ocean

T Rel Inst:

~~<ria~~ <ria <ria cháy săm ràp <ka·n <thi· ka·n pen <ria
 boat boat be-used for activity which activity be boat
 <de·n <nay <samùt> NP>S>NP>S>NP
 walk in ocean

T Red N Del:

~~<ria~~ <ria <ria cháy săm ràp <ka·n <thi· pen <de·n <nay
 boat boat be-used for activity which be walk in
 <samùt> NP>S>NP>S>NP
 ocean

T Rel Pro Pre: Applies vacuously.

T Pron Del:

~~<ria~~ <ria <ria cháy săm ràp <ka·n <pen <de·n <nay <samùt> NP>VP>NP>
 boat boat be-used for activity be walk in ocean

T App:

~~<ria~~ <ria <ria cháy săm ràp <ka·n <de·n <nay <samùt> NP>NP>S>NP
 boat boat be-used for activity walk in ocean

T Nom:

~~<ria~~ <ria <ria cháy săm ràp <ka·ndə·n <nay <samùt> N>S>NP
 boat boat be-used for walking in ocean

FOURTH CYCLE:

T Rel Inst:

~~<ria~~ <thi· <ria <ria cháy săm ràp <ka·ndə·n <nay <samùt> N>S>NP
 boat which boat be-used for walking in ocean

T Red N Del:

~~<ria~~ <thi· <ria <ria cháy săm ràp <ka·ndə·n <nay <samùt> N>S>NP
 boat which be-used for walking in ocean

T cháy Del:

~~<ria~~ <thi· <săm ràp <ka·ndə·n <nay <samùt> N>NP
 boat which for walking in ocean

T Pron Del:

~~<ria~~ <săm ràp <ka·ndə·n <nay <samùt> N>NP
 boat for walking in ocean

T sămràp Del:

<rīa <ka·ndə·n nay samūt>_N>_{NP}
 boat walking in ocean

T ka·n Del:

<rīa <də·n nay samūt>_N>_{NP}
 boat walk in ocean

GNC converts this structure to rīade·nnaysamūt. NCD 2, as it now stands, will delete the preposition nay 'in'. Notice that the prediction implicit in the form of NCD 2, that it will not matter whether the preposition is inserted into the verb phrase originally or whether it is placed there by the Adverbial Promotion Transformation, is correct. However, NCD 2 will not now delete the preposition without also deleting the verb. As a result, we need to divide NCD 2 into two rules:

NCD 2 <N V (Prp) N>_N
 1 2 3 4
 → 1 ∅ 3 4

NCD 3 <N (V) Prp N>_N
 obl 1 2 3 4
 → 1 2 ∅ 4

To illustrate the more direct derivation, we will use nākmenpī·n 'sharpshooter'.

BASE:

<nāk <nāk men dūay pī·n>_S>_{NP}
 expert expert be-accurate with gun

FIRST CYCLE: No operations.

SECOND CYCLE:

T Rel Inst:

<nāk <thī· nāk men dūay pī·n>_S>_{NP}
 expert which expert be-accurate with gun

T Red N Del:

<nák <thi· men dūay pī·n>S>NP
expert which be-accurate with gun

T Pron Del:

<nák <men dūay pī·n>VP>NP
expert be-accurate with gun

In this case also, GNC will generate a compound including the preposition. NCD 3 is adequate to make the necessary adjustment to give nákmenpī·n. Since nák is a bound form, GNC must apply, or an ungrammatical noun phrase results.

Examples of Subject-Verb-Adverbial compounds:

dèkde·ntó? 'child waiter'

dèk 'child'

də·n 'to walk'

tó? 'table'

khonde·ntó? 'waiter'

khon 'person'

də·n 'to walk'

tó? 'table'

manútde·ndin 'ordinary human being'

manút 'man'

də·n 'to walk'

din 'earth'

nákrênpi·n 'sharpshooter'

nák 'expert'⁴⁸

mên 'to be accurate'

pī·n 'gun'

riade·nsamùt 'ocean-going steamer'

riá 'boat'

də·n 'to walk'

samùt 'ocean'

síaphà·?òk 'an upper garment which opens down the front'

síá 'upper garment'

phà· 'to split'

?òk 'chest'

yínda·ntó? 'waitress'

yín 'woman'

də·n 'to walk'

tó? 'table'

4.2.1.6 Adverbial as the Second Member. A larger number of compounds are derived from similar sentences, but have only the subject and adverbial preserved. Since we have already accounted for the Subject-Verb-Adverbial type compound, it should be a simple matter to show how the same rules will also generate Subject-Adverbial compounds. While this is generally the case, Subject-Adverbial compounds are unusual in a number of ways. In the first place, the overwhelming majority of examples of Subject-Adverbial compounds are derivable from a sentence of a very definite form, namely:

N { ?a·sǎyyù· } Prep-of-Location N
 dwell
 yù·
 be-located

But there are a few notable exceptions. ríaphíwnám 'surface vessel' (ría 'boat', phíwnám 'surface of the water') seems to be the counterpart of ríade·nsamùt 'ocean-going steamer'. Another compound, menwan 'housefly' (men 'insect', wan 'day') requires derivation from a sentence which includes a time adverbial. It seems necessary, then, to construct a grammar

which is sufficiently comprehensive to generate all kinds of Subject-Adverbial compounds. Finally, a few compounds of this type retain their prepositions. An example is phûak-tâ·y·din 'underground organization' (phûak 'group', tâ·y 'under', din 'earth'). However, there are no compounds which preserve both the verb and preposition, so we can account for this type of compound if we allow NCD 2 and NCD 3 to be ordered as they are, but restrict NCD 3 a little more carefully.

NCD 3	<N	(V)	Prp	N>	N	Obligatory if 2 is present; optional otherwise.
	1	2	3	4		
→	1	2	∅	4		

Subject-Adverbial compounds appear to the English speaker to be unamenable to cháy sǎmráp derivation, for the most part at least. But when I asked for a definition of thahǎ·nrǎa 'sailor' (thahǎ·n 'soldier', rǎa 'boat'), Mr. Sirisamphan offered both of the following sentences:⁴⁹

thahǎ·n sǎmráp yù· nay rǎa
soldier for be-located in boat

thahǎ·n thǎi· yù· nay rǎa
soldier who be-located in boat

This pair of definitions not only demonstrates the validity of cháy sǎmráp derivation of Subject-Adverbial compounds, but also represents strong evidence for the validity of allowing our grammar to generate compounds which are ambiguous in this way in general. We can use thahǎ·nrǎa as the example for both cháy sǎmráp and direct relative clause derivations. The cháy sǎmráp derivation starts with the base below.

BASE:

<thahă·n <khon chây sāmrap <ka·n <ka·n pen <<thahă·n
 soldier person use for activity activity be soldier
 yù· nay rīa >S>NP>S>NP thahă·n>S>NP
 be-located in boat soldier

A derivation analogous to that of riade·nsamūt 'ocean-going steamer' gives thahă·nyù·nayrīa. The operation of both NCD 2 and NCD 3 adjusts this to thahă·nrīa.

The base for the other alternative is:

BASE:

<thahă·n <thahă·n yù· nay rīa >S>NP
 soldier soldier be-located in boat

A derivation like that for nāmkmnpī·n 'sharpshooter' yields the compound form thahă·nyù·nayrīa with the usual slightly different structure from that of its chây sāmrap counterpart. The same type of application of NCD 2 and NCD 3 generates thahă·nrīa from this structure.

Examples of Subject-Adverbial compounds are these:

thahă·nrīa 'sailor'

thahă·n 'soldier'

rīa 'boat'

khī·fan 'decaying food particles between the teeth'

khī· 'excrement'

fan 'tooth'

cha·wthay 'the Thai'

cha·w 'dweller in'

thay 'Thai'

kây·fá· 'pheasant'

kây 'chicken'

fá· 'sky'

pha·sǎ·nansǎi 'literary language'

pha·sǎ 'language'

nansǎi 'book'

menwan 'housefly'

men 'insect'

wan 'day'

riaphiwnǎ·m 'surface vessel'

riá 'ship, boat'

phiwnǎ·m 'surface of the water'

mǔ·kho·rá·t 'fool'

mǔ 'pig'

kho·rá·t 'Name of a township in Northeastern Thailand'

lèknay 'stinger of an insect'

lèk 'iron'

nay 'in'

cha·wtà·nprathê·t 'foreigner'

cha·w 'dweller in'⁵⁰

tà·nprathê·t 'another country'

4.2.1.7 Special Problems. Among the compounds with the subject as head, there are two seemingly unrelated problems whose solution will prove to be not only related, but of considerable generality throughout the grammar of Thai noun compounds. The first of these has to do with plant names. The word for 'tree' is tôn·mǎ·y (tôn 'stalk', mǎ·y 'wood, stick'). Particular plants are named with a compound consisting of tôn and a word to indicate the type, for example tônklûay 'banana tree' (klûay 'banana') and tônkulà·p 'rose bush' (kulà·p 'rose'). Similarly, the general word for 'flower' is dò·kmá·y (dò·k 'blossom'). Particular flowers are named by compounds with dò·k, e.g.,

dò·kkulà·p 'rose flower' (kulà·p 'rose'). The words involving leaves and fruits are similarly constructed; phõnlamá·y 'fruit' (phõn 'fruit') and baymà·y 'leaf' (bay 'leaf'). The words for particular leaves and fruits are compounds of phõn(la) and bay. The key to these plant names seems to be the successful analysis of tõnmá·y. The definition of phõnlamá·y given me by Mr. Sirisamphan was:

phõn thî· kə·t cà·k tõnmá·y
fruit which originate from tree

"Fruit which comes out of a tree."

The definitions of dò·kmá·y 'flower' and baymá·y 'leaf' followed the same pattern. The definition for dò·kkulà·p 'rose' was:

dò·k thî· ?ò·k cà·k tõnkulà·p
flower which emerge from rose-bush

The definition for tõnkulà·p 'rose bush' was, somewhat circularly:

tõnmá·y thî· mî· dò·kkulà·p
tree which have rose-blossom

The general pattern of all of these definitions fits compound formation rules which are already in the grammar. The first two fit the requirements for Subject-Adverbial compounds and the third for Subject-Object compounds. But for tõnmá·y, Mr. Sirisamphan was unable to give a definition of this type, insisting that it was one word. This response could be accounted for if we recognize tõn as a bound form, in this meaning at least. If so, we could allow tõnmá·y to be generated from some sentence which means something like "A stalk which is wood" or "A plant which resembles a stick". The member tõn, like nák 'expert' and cha·w 'dweller in', would be specified in such

a way that GNC has to apply to any sentence which contains it. This would explain why the informant was not able to give the definition. If this is done, all these plant names could be generated with existing rules with one difficulty.

It soon becomes apparent that the rules we now have will generate compounds like *dò·ktôn·má·y for dò·kmá·y 'flower', *baytôn·má·y for bay·má·y 'leaf', *baytôn·maphra·w for bay·ma·phra·w 'leaf of a coconut tree' and *tôn·dò·kkulà·p for tôn·kulà·p 'rose bush'. An additional rule to delete the unwanted constituent can be proposed in the following form:

$$\begin{array}{rcl} \text{NCD 4} & <\text{N} < \begin{array}{c} +\text{N} \\ +(3)\text{Nu} \end{array} \text{N} > \text{N} > \text{N} \\ \text{obl} \rightarrow & \begin{array}{ccc} 1 & 2 & 3 \\ 1 & \emptyset & 3 \end{array} \end{array}$$

NCD 4 obligatorily deletes one constituent of a noun compound if 1) it is also a constituent of a compound which in turn has been compounded with another noun and 2) it is a noun with the feature $[(+)(3)\text{Nu}]$, that is, it is a classifier for the noun with which it is most intimately compounded. The nouns tôn, dò·k, bay, and phôn are classifiers for plants, flowers, leaves, and fruits, respectively.

The other problem has to do with certain nouns of location. Some of these consist of a noun referring to a section of time or space and a preposition or noun of location. Some of these are khâ·nnay 'inside' (khâ·n 'side', nay 'in'), bīanlăn 'behind, in the past' (bīan 'part', lăn 'behind') and pha·ytâ·y 'underneath, underside' (pha·y 'part', tâ·y 'under'). In addition,

there are similar compounds with other nouns as heads, such as khonnay 'insider' (khon 'person', nay 'in') and khriannay 'internal body organs' (khriān 'material', nay 'in'). The solution that would do the least violence to the rules we already have would be to derive these as Subject-Adverbial compounds of which the object of the locative preposition is a Pre-noun deleted by NCD 1 and the verb is deleted by NCD 2, leaving the subject and preposition. Thus, khâ·nnay 'inside' would have as a base:

$\langle \text{khâ} \cdot \text{n} \quad \langle \text{khâ} \cdot \text{n} \quad \text{yù} \cdot \quad \text{nay} \quad \text{khō} \cdot \text{n} \rangle \text{S} \rangle \text{NP}$
 side side be-located in thing

If this is the case, it would be expected that a Thai speaker would accept khâ·n thī· yù· nay khō·n 'side which is in something' as a definition for khâ·nnay. This was suggested to two informants independently and both rejected the suggestion firmly. When asked to define words like biannay 'inside' and khonnay 'insider', or similar words, both independently suggested:

$\frac{\text{bīan}}{\text{part}} \quad \frac{(\text{khon})}{(\text{person})} \quad \frac{\text{thī} \cdot \quad \text{yù} \cdot}{\text{which be-located}} \quad \frac{\text{khâ} \cdot \text{nnay}}{\text{inside}}$

If we take these definitions seriously, it leads us to suspect that the khâ·n compounds occupy the crucial position which tôn·má·y occupies relative to plant-name compounds. If this is so, the same solution is open. The word khâ·n 'side' is marked as bound, so that the base structure above must be carried through GNC to yield a grammatical form.⁵¹

This solution leaves us with compounds like *bīankhâ·nnay

and *khonkhâ·nnay. Clearly, NCD 4 will not work to correct these compounds since prepositions do not take classifiers. But perhaps another specification which is more inclusive than the classifier specification can handle the problem. The revised form of NCD 4 would be something like:

$$\begin{array}{c} \text{NCD 4} \quad <N < \left[\begin{array}{c} +N \\ +\text{Gen w/r (3)} \end{array} \right] \left[\begin{array}{c} +N \\ +\text{Spc w/r (2)} \end{array} \right] >N >N \\ \text{obl} \rightarrow \begin{array}{ccc} 1 & 2 & 3 \\ 1 & \emptyset & 3 \end{array} \end{array}$$

The features "Generic with respect to 3" and "Specific with respect to 2" are approximations of a sort of concord restriction between the meanings of the forms designated 2 and 3. Thus the meaning of nay 'in' is "specific" with regard to an area of meaning with regard to which the meaning of khâ·n 'side' is "generic". A similar relationship holds between dò·k 'blossom' and kulâ·p 'rose'.

NCD 4 is in obvious need of a great deal of refinement if it is to have a place in Thai grammar. However, it seems to be an approximation of a principle which could have a great deal to do with the grammar of Thai noun compounds. In the course of our discussion of Thai compounds so far, we have glossed over some facts which could have a bearing on the matters under discussion. In giving the definition of mũ·să·mchán 'bacon' (mũ· 'pork', să·m 'three', chán 'layer'), Mr. Sirisamphan did not give:

mũ· thi· mi· să·m chán
pork which have three layer
"Pork which has three layers"

but rather:

mũ· suan thĩ· mĩ· să·m chăn
pork part which have three layer

"The part of pork which has three layers"

Similarly, the definition of wuanom 'milch cow' given in the Thai-Thai Dictionary⁵² is not:

wua thĩ· hây nom mă·k
cow which give milk be-much

"A cow which gives a lot of milk"

but rather:

wua phan thĩ· hây nom mă·k
cow kind which give milk be-much

"A kind of cow which gives a lot of milk"

In addition, there are a few compounds which seem to involve meanings like "part". For example, the word for 'egg yolk' is khâyde·n (khây 'egg', de·n 'to be red') and ta·dam means 'pupil' (ta· 'eye', dam 'to be black'). It seems clear that the origins of these compounds should be something like:

N suan thĩ· V color
part which

"The part of N which is a certain color"

If we allow a classifier like phan 'kind' to appear in the source sentences of our compounds, it could provide part of the answer to the problem posed in connection with Miss Udomphol's frame for testing separability in compounds with adjectival verbs. If námhổ·m 'perfume' (nám 'fluid', hổ·m 'to be fragrant') is derived from the structure which underlies:

nám phan thĩ· hổ·m
fluid kind which be-fragrant

"A kind of fluid which is fragrant"

rather than:

nám thi• hố•m
fluid which be-fragrant
"Fluid which is fragrant"

Then the sentence:

námhố•m khàt ní• mây hố•m lè•y
perfume bottle this not be-fragrant at-all
"This bottle of perfume (lit. fragrant fluid) is not
fragrant at all."

does not seem contradictory.

If classifiers should appear in derivations in this manner, GNC would apparently produce structures like *námphanhố•m and *khàysuandè•n. If the grammar of classifiers in Thai is such that nám and phan are immediate constituents and námphan and hố•m are immediate constituents at another level, then there seems to be a general principle of compound deletion which handles these cases as well as those like tônmá•y 'tree' and khà•nnay 'inside'. The principle seems to be something like the following:

Principle: If two immediate constituents XY are compounded with a third constituent Z, and if X has a generic meaning in an area in which Y has a specific meaning, then delete X; or vice versa.

This notion is sufficiently vague and the grammar of classifiers sufficiently unclear as to place further pursuit of the problem beyond the range of this study. We will, however, have occasion to refer to it in connection with another problem.

Our examples of plant-name compounds are:

dò·kklûaymá·y 'orchid'

dò·k 'blossom'

(tôn)klûaymá·y 'orchid plant'

dò·kkulà·p 'rose'

dò·k 'blossom'

(tôn)kulà·p 'rose bush'

dò·kmá·y 'flower'

dò·k 'blossom'

(tôn)má·y 'tree, plant'

tônklûay 'banana tree'

tôn 'plant'

klûay 'banana'

tônkulà·p 'rose bush'

tôn 'plant'

(dò·k)kulà·p 'rose'

phõnlamá·y 'fruit'

phõn(la) 'fruit'

(tôn)má·y 'tree'

baymaphrá·w 'leaf of the coconut tree'

bay 'leaf'

(tôn)maphrá·w 'coconut tree'

baymá·y 'leaf'

bay 'leaf'

(tôn)má·y 'tree'

Our examples of Subject-Adverbial compounds whose second members are prepositions or locative nouns are these:

ka·nke·nnay 'undershorts, panties'

ka·nke·n 'pants'

nay 'in'

khâ·nnô·k 'outside, the outside'

khâ·n 'side'

nô·k 'outer'

khâ·nnâ 'in front, ahead'

khâ·n 'side'

nâ 'front'

khonnô·k 'outsider'

khon 'person'

nô·k 'outside'

khriannay 'internal body organs'

khriān 'material'

nay 'in'

chānnay 'inner layer'

chān 'layer'

nay 'in'

pha·ynay 'inside, internal'

pha·y 'part, side'

nay 'in'

chānlān 'the later generation'

chān 'layer'

lān 'behind'

bā·nnô·k 'the country'

bā·n 'village'

nô·k 'outer'

biānbon 'upper part'

biān 'Classifier for parts'

bon 'upper, on'

4.2.2 Compounds with Object as Head. Before going into detail concerning compounds which preserve the object of an underlying sentence as head, we should consider the potential impact of the Ambivalent Verb Transformation on the generation of such compounds. If it could be shown that all Thai verbs are ambivalent, there would be no need to set up a special mechanism for the derivation of verbs with the object as head,

except in cases in which the subject is also preserved. For the rest, T Am Vb could apply, placing the object in subject position and deleting the Pro-noun subject. The compounds could then be generated exactly like Subject-Verb or Subject-Adverbial compounds. To obtain khǒ·nkin 'edibles' (khǒ·n 'thing', kin 'to eat'), T Am Vb would convert:

khon kin khǎ·n
person eat thing
"People eat things."

into:

khǎ·ŋ kin
thing be-eaten
"Things are eaten."

From this point on, the derivation could follow the lines of derivation of Subject-Verb compounds like má·khè·ŋ 'race horse' (má· 'horse', khè·ŋ 'to compete') or riabin 'airplane' (ri 'boat', bin 'to fly'). This is not done for two reasons: 1) not all the verbs involved are ambivalent, and 2) it would not work for Object-Subject compounds anyway.⁵³

4.2.2.1 Verb as Second Member. Apparently there are no three-member compounds whose heads are the objects of embedded sentences, so we will first consider the Object-Verb type.⁵⁴ The compound rót·do·ysă·n 'bus' (rót 'bus', do·ysă·n 'to travel by') illustrates the cháy sāmrap derivation.

BASE:

<rót <khon cháy săm rập <ka·n <ka·n pen <<khon
vehicle person use for activity activity be person

do·vsà·n rót >S>NP>S>NP rót
travel-by vehicle vehicle >S>NP

FIRST CYCLE: No operations.

SECOND CYCLE: No operations.

THIRD CYCLE:

T Am Vb:

<rót <rót cháy săm ràp <ka·n <ka·n pen <khon
vehicle vehicle be-used for activity activity be person
do·ysă·n rót >NP>S>NP>S>NP
travel-by vehicle

T Rel Inst:

<rót <rót cháy săm ràp <ka·n <thi· ka·n pen
vehicle vehicle be-used for activity which activity be
<khon do·ysă·n rót >NP>S>NP>S>NP
person travel-by vehicle

T Red Nom Del:

<rót <rót cháy săm ràp <ka·n <thi· pen <khon
vehicle vehicle be-used for activity which be person
do·ysă·n >NP>S>NP>S>NP
travel-by

T Rel Pro Pre: Applies vacuously.

T Pron Del:

<rót <rót cháy săm ràp <ka·n <pen <khon
vehicle vehicle be-used for activity be person
do·ysă·n >NP>VP>NP>S>NP
travel-by

T App:

<rót <rót cháy săm ràp <ka·n <khon
vehicle vehicle be-used for activity person
do·ysă·n >NP>NP>S>NP
travel-by

T Nom:

<rót <rót cháy săm ràp <ka·ndo·ysă·n khố·n khon >NP>S>NP
vehicle vehicle be-used for traveling-by of person

T Pro Del:

<rót <rót cháy săm ràp <ka·ndo·ysă·n khỗ·n>NP>S>NP
vehicle vehicle be-used for traveling-by of

T Is Prep Del:

<rót <rót cháy săm ràp <ka·ndo·ysă·n>N>S>NP
vehicle vehicle be-used for traveling-by

FOURTH CYCLE:

T Rel Inst:

<rót <thi· rót cháy săm ràp <ka·ndo·ysă·n>N>S>NP
vehicle which vehicle be-used for traveling-by

T Red N Del:

<rót <thi· cháy săm ràp <ka·ndo·ysă·n>N>S>NP
vehicle which be-used for traveling-by

T cháy Del:

<rót thi· săm ràp <ka·ndo·ysă·n>N>NP
vehicle which for traveling-by

T Pron Del:

<rót săm ràp <ka·ndo·ysă·n>N>NP
vehicle for traveling-by

T săm ràp Del:

<rót <ka·ndo·ysă·n>N>NP
vehicle traveling-by

T ka·n Del:

<rót <do·ysă·n>N>NP
vehicle traveling-by

GNC directly creates rótdo·ysă·n 'bus' from this structure.

The direct relative clause derivation can be illustrated with sàtlían 'domestic animal'.

BASE:

<sàt <khon lían sàt>S>NP
animal person care-for animal

FIRST CYCLE: No operations.

SECOND CYCLE:

T Rel Inst:

<sàt <khon lian thi. sàt >_S>_{NP}
 animal person care-for which animal

T Red N Del:

<sàt <khon lian thi. >_S>_{NP}
 animal person care-for which

T Rel Pro Pre:

<sàt <thi. khon lian >_S>_{NP}
 animal which person care-for

T Pron Del:

<sàt <khon lian >_S>_{NP}
 animal person care-for

For GNC to operate on this structure, it will require still another revision. The new rule is:

$$\begin{array}{c} \text{GNC} <N \begin{Bmatrix} N (V) \\ V (\text{Prp}) (N) \\ \underline{mi.} (\text{Nu}) N \end{Bmatrix} >_N \\ \rightarrow \begin{matrix} 1 \\ 1 > 2 \end{matrix} \end{array}$$

Application of GNC in this form produces sàtkhonlian. NCD 1 is designed to delete Pro-nouns in compounds. However, it will not delete khon in this position and will have to be revised. The new NCD 1 is:⁵⁵

$$\begin{array}{c} \text{NCD 1} <(N V) \begin{bmatrix} +N \\ +\text{Pro} \end{bmatrix} (V N) >_N \\ \rightarrow \begin{matrix} 1 & 2 & 3 \\ 1 & \emptyset & 3 \end{matrix} \end{array} \quad \begin{array}{l} \text{Where 2 and 3 are neither} \\ \text{both present nor both ab-} \\ \text{sent.} \end{array}$$

The new versions of GNC and NCD 1 now produce sàtlian 'domestic animal'. It is the case again that the chây sảmràp derivation is apparently simpler in a deep sense than is the apparently

more direct method. Examples of Object-Verb compounds are listed below.

rótdo·ysă·n 'bus'

rót 'vehicle'

do·ysă·n 'to travel by a vehicle'

sàtlían 'domestic animal'

sàt 'animal'

lían 'to care for'

khǒ·nkin 'edibles'

khǒ·n 'thing'

kin 'to eat'

thale·sà·p 'lake'

thale· 'sea'

sà·p 'to curse'

phâ·pkían 'a drawing'

phâ·p 'picture'

khían 'to draw, write'

miakèp 'mistress'

mia 'wife'

kèp 'to keep'

rótlâ·k 'rickshaw'

rót 'vehicle'

lâ·k 'to pull, tow'

ríaphûan 'tugboat, boats pulled by a tugboat'

ría 'boat, ship'

phûan 'to connect, attach'

sâykrò·k 'sausage'

sây 'intestines, inside part'

khò·k 'to fill'

ʔa·hǎ·nka·nkin 'nutriment'

ʔa·hǎ·n 'food'

ka·nkin 'eating'

4.2.2.2 Subject as Second Member. Object-Subject compounds can be divided into two sections. One deals with underlying sentences with mi· 'to have' and results in possessive compounds. Another section deals with compounds from sentences with all other verbs. The possessive compounds do not have cháy sǎmráp derivations, but their derivation is unique. The word for 'fishbone', ka·npla·, will serve as an example.

BASE:

<ka·n <pla· mi· ka·n>S>NP
bone fish have bone

FIRST CYCLE: No operations.

SECOND CYCLE:

T Rel Inst:

<ka·n <pla· mi· thǐ· ka·n>S>NP
bone fish have which bone

T Rel N Del:

<ka·n <pla· mi· thǐ· >S>NP
bone fish have which

T Rel Pro Pre:

<ka·n <thǐ· pla· mi· >S>NP
bone which fish have

T Pron Del:

<ka·n <pla· mi· >S>NP
bone fish have

T Poss:

<ka·ŋ khǒ·ŋ pla·>_{NP}
bone of fish

T khǒ·ŋ Del:

<ka·ŋ pla·>_{NP}
bone fish

With this type of compound all the rearrangement and deletion operations are completed before operation of GNC. GNC converts the above structure into ka·ŋpla· 'fishbone'. A few examples are not quite so simple. prathê·tthay 'Thailand' (prathê·t 'country', thay 'Thai') must be derived via prathê·t thi· khonthay mi· 'The country which the Thai people have' since thay is a bound form. Several other similar compounds must also be derived in this way. The derivation illustrated above in these cases ends in *prathê·tkhonthay. But this is another example like dò·kmá·y 'flower' and khonnay 'insider'. This, then, is another instance in which the noun compound deletion principle sketched above is needed. Examples of possessive compounds are listed below.

ka·ŋpla· 'fishbone'

ka·ŋ 'bone'

pla· 'fish'

khò·pthanǒn 'curb'

khò·p 'edge, rim'

thanǒn 'street'

dâ·mpɿ·n 'gunstock'

dâ·m 'handle'

pɿ·n 'gun'

níwthá·w 'toe'

nfw 'digit'

thá·w 'foot'

fà·thá·w 'sole of the foot'

fà 'palm, sole'

thá·w 'foot'

fà·m4 'palm of the hand'

fà: 'palm, sole'

m4 • 'hand'

phô-phũa 'woman's father-in-law'

phô • 'father'

phũa 'husband'

lû·kpla· 'young fish'

lû·k 'offspring, child'

pla· 'fish'

sâykày 'giblets'

sây 'intestines'

kà'y 'chicken'

hỗ·nkày 'cock's comb'

ṇō·n 'comb of a fowl'

kà:y 'chicken'

The remaining Object-Subject compounds are non-possessive.

The word rótmá 'horse-drawn carriage' seems to be a candidate for cháy sảmráp derivation. But an attempt to derive it by this method will pose serious problems. The derivation would begin with the base below:

BASE:

<rót <khon cháy săm ràp ka·n <ka·n pen <<má·
vehicle person use for activity activity be horse

$\frac{\text{l\`a} \cdot \text{k}}{\text{pull}} \frac{\text{r\`ot}}{\text{vehicle}} >_{\text{S}} >_{\text{NP}} >_{\text{S}} >_{\text{NP}} \frac{\text{r\`ot}}{\text{vehicle}} >_{\text{S}} >_{\text{NP}}$

The derivation follows that of rótđo·ysă·n 'bus' up to T Nom.

The structure preceding T Nom is:

<rót <rót cháy sămráp <ka·n <má· lâ·k>NP>NP>S>NP
vehicle vehicle be-used for activity horse pull

After application of T Nom, the structure is:

<rót <rót cháy sămráp <ka·nlâ·k khố·n má·>NP>S>NP
vehicle vehicle be-used for pulling of horse

T Pro Del will not operate on this structure and as a result,

the conditions for application of T Is Prep Del are not met.

Thus, the derivation ends in:

<rót <lâ·k khố·n má·>NP>NP
vehicle pulling of horse

GNC clearly does not operate on such a structure. Three solutions are possible. GNC could be revised to operate on this structure and post-compounding transformation operations could be set up to give the correct order. But this solution would be woefully ad hoc and non-revealing. The second possibility would be to claim that no Object-Subject compounds are derived by cháy sămráp derivation at all. There seems to be some evidence for this, since rótmá· need not be so derived and it appears that relatively few compounds of this type are open to this derivation. The third solution would be to allow some compounds to be derived from still another base, namely:

<rót <khon cháy sămráp <<má· lâ·k rót>S>NP
vehicle person use for horse pull vehicle
rót >S>NP
vehicle

The nominalization step would be eliminated in these cases and

as a result the unwanted transposition of the subject would not occur. The derivation would end in:

<rót <má· lâ·k> NP>NP
vehicle horse pull

This can be readily handled by GNC and NCD 2 to produce rótmá·.

Another result of this proposal would be the elimination of T ka·n Del from the derivation of noun compounds (although not from the grammar, since it is independently motivated). Compounds which preserve ka·n would be derived via nominalization and those which do not from the above structure in whose derivation ka·n never appears. The disadvantage is that this solution introduces a third way in which noun compounds without verbs as a member can be generally ambiguous; and it is questionable whether or not this is justifiable. Our tentative solution is the second one, in which we assume there is no cháy sămràp derivation for these compounds.

Another word for a kind of vehicle, rótcèk 'rickshaw', can be derived by the direct relative clause process.

BASE:

<rót <khoncèk cháy rót >S>NP
vehicle Chinese use vehicle

A derivation process similar to that of sàtlían 'domestic animal' produces rótkhoncèkcháy after application of GNC. The verb cháy 'to use' is deleted by NCD 2 and the Pro-form khon by the rule which expresses the noun compound deletion principle.

Examples of Object-Subject compounds which are not possessive are:

rótcèk 'rickshaw'

rót 'vehicle'

(khon)cèk 'Chinese people'

rótmá 'horse-drawn carriage'

rót 'vehicle'

má 'horse'

sà·sană·khrít 'Christianity'

sà·sană 'religion'

khrít 'Christ' (from European)

sà·sană·phút 'Buddhism'

sà·sană 'religion'

phút 'Buddha'

khí·kòp 'wood shavings'

khí 'excrement'

kòp 'carpenter's plane'

khí·kày 'chicken droppings'

khí 'excrement'

kày 'chicken'

khí·thâw 'ashes'

khí 'excrement'

thâw 'stove'

thanõmlúan 'public road'

thanõm 'street'

lúan 'government'

thunrátthaba·n 'government scholarship'

thun 'capital'

rátthaba·n 'government'

bòtkawī 'poetry'

bòt 'poem in a certain verse form'

kawī 'poet'

4.2.2.3 Adverbial as Second Member. The third and last kind of compound which has an object head is the Object-Adverbial type. Of particular interest here is a large set of compounds which appear to be derived from a sentence whose predicate has tham dūay 'to make with' or 'to make out of'.⁵⁶ No new derivation processes are needed for this set since in this instance early application of the Ambivalent Verb Transformation allows the original object to be compounded as a subject. Thus kronlèk 'iron cage' is derived from the base:

BASE:

<kron <khon tham kron dūay lèk >_S>_{NP}
 cage person make cage with iron

But immediate application of T Am Vb converts this to:

<kron <kron tham dūay lèk >_S>_{NP}
 cage cage be-made with iron

From this structure, kronlèk can be derived exactly like a Subject-Adverbial compound like thahă·nr̥a 'sailor'. For these compounds, the cháy sāmrap derivation is not possible. The list of tham dūay compounds is:

kronlèk 'iron cage'

kron 'cage'

lèk 'iron'

nūnluat 'screen'

nūn 'mosquito net'

luat 'wire'

kradà·tdi·òuk 'tin foil'

kradà·t 'paper'

di·bùk 'tin'

tû·kracòk 'glass cabinet'

tû· 'cabinet'

kracòk 'glass'

ke·nkày 'chicken curry'

ke·n 'curry'

kày 'chicken'

fo·nná·m 'bubbles, foam'

fo·n 'bubbles, foam'

ná·m 'water'

riánmá·y 'wooden building'

rián 'house'

má·y 'wood'

săwkhò·nkhri·t 'concrete post'

săw 'post, pole, pillar'

khò·nkhri·t 'concrete' (from English)

thà·nmá·y 'charcoal from wood'

thà·n 'charcoal'

má·y 'wood'

thàle·sa·y 'desert'

thàle· 'sea'

sa·y 'sand'

There are also some other compounds of the Object-Adverbial type. Of these, khriánkhrua 'kitchen utensils' is one which can be derived via cháy sămràp.

BASE:

<khrián <khon cháy sămràp <ka·n <ka·n pen <<khon
implements person use for activity activity be person

cháy khrián nay khrua >S>NP>S>NP khrián >S>NP ⁵⁷
use implement in kitchen implement

The derivation continues parallel to that of rótđo·ysă·n 'bus' except that the adverbial nay khrua is carried along throughout. The pre-GNC structure is :

<khriān <khon cháy nay khrua >N>NP
 implement person use in kitchen

GNC makes this khriānkhoncháyaykhrua. To get khriānkhrua, the three Noun Compound Deletion transformations must all apply. NCD 1 deletes khon, NCD 2 eliminates cháy, and NCD 3 deletes nay. Another word referring to the kitchen, ka·nkhrua 'cooking', will illustrate the direct relative clause derivation process:

BASE:

<ka·n <khon tham ka·n nay khrua >S>NP
 activity person do activity in kitchen

The derivation follows that of sătliān 'domestic animal', except that the adverbial is carried through the derivation. The structure of the last line is:

<ka·n <khon tham nay khrua >S>NP
 activity person do in kitchen

There is no form of GNC which will create a compound noun out of this structure. It is now necessary to state GNC in two cases. Case a) generates compounds from structures which already consist of only two nouns. Case b) generates compounds from structures containing various deformations of relative clauses.

GNC <N { a) N
 { (N) (V(Prp))
 b) < { (V(Prp)) (N)
 m1. (Nu) N } >X } >NP
 2
 1
 → 1>2

Case a) must be separated from case b), even when X is N. When the pre-GNC structure consists of nouns only, the second noun will not have the structure $\langle N \rangle_N$ since the upper N would be deleted by the node-labeling principles. This means that cháy săm rập compounds are generated by case b) of GNC where X is N.

This form of GNC will generate ka·nkhonthamnaykhrua, which is adjusted by application of NCD 1-3 to give ka·nkhrua 'cooking'. Examples of Object-Adverbial compounds are in the list below.

ʔahă·nchăw 'breakfast'

ʔahă·n 'food'

(tə·n)chăw 'morning'⁵⁸

khriankhrua 'kitchen utensils'

khriān 'implement'

khrua 'kitchen'

khāypà· 'malaria'

khāy 'sickness'

pà· 'forest'

ka·nkhrua 'cooking'

ka·n 'activity'

khrua 'kitchen'

dinpɛ̄·n 'gunpowder'

din 'explosive powder'

pɛ̄·n 'gun'

tūmkhũ· 'earring'

tūm 'suspended object'

hũ· 'ear'

pha·sí·nendā·y 'income tax'

pha·sí· 'tax'

nendā·y 'income'

mí·cháu 'breakfast'

mí· 'Classifier for meals'

(tə·n)cháu 'morning'

riapha·y 'boat propelled by paddling'

ria 'boat'

pha·y 'paddle'

?ahă·nkla·nwan 'lunch'

?ahă·n 'food'

(tə·n)kla·nwan 'mid-day'

4.2.3 Compounds with Adverbial as Head.

4.2.3.1 Subject and Verb as Second and Third Members.

Of compounds with adverbials as the first member, there are a few examples in my collection of which the second and third members are the underlying subject and verb of the innermost embedded sentence. In none of these examples is the cháy sāmrap derivation appropriate. The relative clause derivation for rídu·baymá·yrúan 'autumn' serves as an example of this type of derivation.

BASE:

<rídu· <baymá·y rúan nay rídu· >_S>_{NP}
 season leaf fall in season

FIRST CYCLE: No operations.

SECOND CYCLE:

T Rel Inst:

<rídu· <baymá·y rúan nay thi· rídu· >_S>_{NP}
 season leaf fall in which season

T Rel N Del:

<rídu· <baymá·y rúan nay thi· >_S>_{NP}
 season leaf fall in which

T Rel Pro Pre:

<rídu· <thi· baymá·y rúan nay>S>NP
 season which leaf fall in

T Pron Del:

<rídu· <baymá·y rúan nay>S>NP
 season leaf fall in

T Is Prep Del:

<rídu· <baymá·y rúan>S>NP
 season leaf fall

GNC generates rídu·baymá·y 'autumn' directly under case b).

Examples of Adverbial-Subject-Verb compounds are:

rídu·baymá·yphlí? 'spring'

rídu 'season'

baymá·y 'leaf'

phlí? 'to bud'

rídu·baymá·yrúan 'autumn'

rídu 'season'

baymá·y 'leaf'

rúan 'to fall'

thíttawantòk 'West'

thít 'direction'

tawan 'sun'

tòk 'to drop'

thíttawan?ò·k 'East'

thít 'direction'

tawan 'sun'

?ò·k 'to emerge'

4.2.3.2 Verb and Object as Second and Third Members.

More common are three-member adverbial-head compounds in which the second and third members are the underlying verb and object. The compound rá·nkhă·ykafə 'coffee shop' is a compound

of this composition which is derivable by the cháy săm ràp derivation.

BASE:

<rá·n <khon cháy rá·n săm ràp <ka·n <ka·n pen
shop person use shop for activity activity be
<khon khă·y kafê· nay rá·n >S>NP>S>NP rá·n >S>NP
person sell coffee in shop shop

FIRST CYCLE: No operations.⁵⁹

SECOND CYCLE: No operations.

THIRD CYCLE:

T Am Vb:

<rá·n <rá·n cháy săm ràp <ka·n <ka·n pen <khon
shop shop be-used for activity activity be person
khă·y kafê· nay rá·n >NP>S>NP>S>NP
sell coffee in shop

T Rel Inst:

<rá·n <rá·n cháy săm ràp <ka·n <thi· ka·n pen
shop shop be-used for activity which activity be
<khon khă·y kafê· nay rá·n >NP>S>NP>S>NP
person sell coffee in shop

T Red N Del:

<rá·n <rá·n cháy săm ràp <ka·n <thi· pen <khon
shop shop be-used for activity which be person
khă·y kafê· nay >NP>S>NP>S>NP
sell coffee in

T Rel Pro Pre: Applies vacuously.

T Pron Del:

<rá·n <rá·n cháy săm ràp <ka·n <pen <khon khă·y
shop shop be-used for activity be person sell
kafê· nay >NP>VP>NP>S>NP
coffee in

T App:

<rá·n <rá·n cháy săm ràp <ka·n <khon khă·y kafê·
shop shop be-used for activity person sell coffee

nay >NP>NP>S>NP
in

T Nom:

<rá·n <rá·n cháy săm ràp <ka·nkhă·y kafê· nay khổ·n
shop shop be-used for selling coffee in of

khon >NP>S>NP
person

T Pro Del:

<rá·n <rá·n cháy săm ràp <ka·nkhă·y kafê· nay khổ·n >NP>S>NP
shop shop be-used for selling coffee in of

T Is Prep Del:

<rá·n <rá·n cháy săm ràp <ka·nkhă·y kafê· >N>S>NP
shop shop be-used for selling coffee

FOURTH CYCLE:

T Rel Inst:

<rá·n <thi· rá·n cháy săm ràp <ka·nkhă·y kafê· >N>S>NP
shop which shop be-used for selling coffee

T Red N Del:

<rá·n <thi· cháy săm ràp <ka·nkhă·y kafê· >N>S>NP
shop which be-used for selling coffee

T cháy Del:

<rá·n thi· săm ràp <ka·nkhă·y kafê· >N>NP
shop which for selling coffee

T Pron Del:

<rá·n săm ràp <ka·nkhă·y kafê· >N>NP
shop for selling coffee

T săm ràp Del:

<rá·n <ka·nkhă·y kafê· >N>NP
shop selling coffee

T ka·n Del:

<rá·n <khă·y kafē· >_N>_{NP}
shop selling coffee

This structure is easily converted by case b) of GNC into rá·nkhă·ykafē· 'coffee shop'. There are no convincing examples of Adverbial-Verb-Object examples which can be derived only by the direct relative clause method, but any example clearly can be so derived in our grammar. To illustrate this type of derivation, we show the process as it operates on ro·n-kèprót 'garage'.

BASE:

<ro·n <khon kèp rót nay ro·n >_S>_{NP}
building person store vehicle in building

FIRST CYCLE: No operations.

SECOND CYCLE:

T Rel Inst:

<ro·n <khon kèp rót nay thi· ro·n >_S>_{NP}
building person store vehicle in which building

T Red N Del:

<ro·n <khon kèp rót nay thi· >_S>_{NP}
building person store vehicle in which

T Rel Pro Pre:

<ro·n <thi· khon kèp rót nay >_S>_{NP}
building which person store vehicle in

T Pron Del:

<ro·n <khon kèp rót nay >_S>_{NP}
building person store vehicle in

T Is Prep Del:

<ro·n <khon kèp rót >_S>_{NP}
building person store vehicle

Case b) of GNC generates ro·nkhonkèprót. NCD 1 deletes khon.

Examples of Adverbial-Verb-Object compounds are:

còtmă·yné?na·mtua 'letter of recommendation'

còtmă·y 'letter'

né?na·m 'to recommend'

tua 'self'

ră·nkhă·ykafe· 'coffee shop'

ră·n 'shop'

khă·y 'to sell'

kafe· 'coffee' (from European)

thi·khiaburi· 'ashtray'

thi· 'place'

khiā 'to scratch as a chicken does'

buri· 'cigarette'

baybè·knen 'withdrawal slip'

bay 'Classifier for leaves and sheets of paper'

bè·k 'to withdraw'

nen 'money'

phā·chétnā· 'handkerchief'

phā· 'cloth'

chét 'to wipe'

nā· 'face'

rótdàpphlə·n 'fire engine'

rót 'vehicle'

dàp 'to extinguish'

phlə·n 'fire'

ro·nkhèprót 'garage'

ro·n 'building'

kèp 'to keep, store'

rót 'vehicle'

rômchu·chî·p 'parachute'

rôm 'umbrella'

chu· 'to save'

chî·p 'life'

hínlápmi·t 'whetstone'

hín 'rock'

láp 'to sharpen'

mi·t 'knife'

?ù·tò·rɿa 'shipyard'

?ù· 'place where something is harbored, cradled, or stored'

tò· 'to construct'

rɿa 'boat, ship'

4.2.2.3 Subject as the Second Member. Other compounds with adverbial heads preserve the subject as the second member. As has been discussed in connection with Object-Subject compounds, cháy sǎmràp derivation causes severe difficulties for the generation of compounds which preserve the underlying subject. Again, we notice that there are few Adverbial-Subject compounds which seem amenable to cháy sǎmràp derivation and none that demand it. We have already seen that cháy sǎmràp derivation does not apply to the few examples of Adverbial-Subject-Verb compounds. We will adhere to our tentative solution which is to assume that no compounds with subjects as second members take cháy sǎmràp derivations.

krabò·kta· 'eye socket' is a compound which illustrates the direct derivation process.

BASE:

<krabò·k <ta· yù· nay krabò·k >_S>_{NP}
cylinder eye be-located in cylinder

The derivation follows the pattern of the ro·nkèprót 'garage' derivation, except that yù· 'to be located' does not have an object. The transformations up to GNC give:

<krabò·k <ta· yù· >_S>_{NP}
 cylinder eye be-located

GNC and NCD 2 yield krabò·kta· 'eye socket'. Examples of this type of compound are:

krabò·kta· 'eye socket'

krabò·k 'cylinder'

ta· 'eye'

khùmkhôn 'hair follicle'

khùm 'pit, cavity'

khôn 'hair'

thă·ntháp 'military base'

thă·n 'base'

tháp 'troop, army'

tó?khríanpê·n 'dresser'

tó? 'table'

khríanpê·n 'powder'

tha·ncara·cò·n 'traffic lane'

tha·n 'way'

cara·cò·n 'traffic'

bò·nen 'silver mine'

bò· 'pit, well'

nen 'silver'

míanci·n 'China, Chinatown'

mían 'city, country'

(khon)ci·n 'Chinese people'

randum 'buttonhole'

ran 'nest'

dum 'button'

radàpná·m 'water level'

radàp 'level'

ná·m 'water'

rídu·fõn 'rainy season'

rídu· 'season'

fõn 'rain'

4.2.2.4 Verb as Second Member. Other adverbial-head compounds have only the verb as second member. There are two kinds of compounds of this form: Type a) compounds are exactly analogous to Adverbial-Verb-Object compounds except the underlying object is a Pro-noun which is deleted. Type b) compounds contain intransitive verbs. An example of a type a) Adverbial-Verb compound derivable via cháy sảmráp is rá·nkhá· 'store'.

BASE:

<rá·n <khon cháy sảmráp <ka·n <ka·n pen <<khon khá·
shop person use for activity activity be person trade

khỏ·n nay rá·n >S>NP>S>NP rá·n >S>NP
thing in shop shop shop

The derivation is the same as that of rá·nkhă·ykafə· 'coffee shop' except that the Pro-noun object khỏ·n 'thing' is deleted by NCD 1. Similarly, ro·nkèp 'storehouse' can be derived from:

BASE:

<ro·n <khon kèp khỏ·n nay ro·n >S>NP
building person store thing in building

in the same way as ro·nkèprót 'garage' is derived, except that khỏ·n is deleted. Examples of type a) Adverbial-Verb compounds are:

khriankhian 'stationery'

khriank 'implement'

khian 'to write'

kradà·tkópi 'carbon paper'

kradà·t 'paper'

kópi 'to copy' (from English)

khamche·n 'invitation'

kham 'word'

che·n 'to invite'

rá·nkhá 'store'

rá·n 'store'

khá 'to trade'

ro·nkep 'storehouse'

ro·n 'building'

kep 'to keep, store'

să·yyùt 'stop wire on a bus'

să·y 'line'

yùt 'to stop'

hônrian 'classroom'

hôn 'room'

rian 'to study'

hũ·rũ·t 'zipper tab'

hũ 'ear'

rũ·t 'to slide, pull'

thĩ·khẽ·n 'pen, hook'

thĩ 'place'

khẽ·n 'to hang'

bayphát 'fan'

bay 'Classifier for sheet-like objects'

phát 'to fan'

Type b) Adverbial-Verb compounds are derived like type a) compounds except that there are no objects with the intransitive verbs and no need for NCD 2. Examples of type b) compounds are listed below.

cùtdiat 'boiling point'

cùt 'spot, point'

diat 'to boil'

tha.nkhin 'runway'

tha.n 'way'

khin 'to rise'

thi.thamka.n 'office'

thi. 'place'

thamka.n 'to do work'

ya.salòp 'anesthetic'

ya. 'medicine'

salòp 'to lose consciousness'

ro.nsuat 'church'

ro.n 'building'

suat 'to chant, recite prayers'

wanchalẽ.m 'day of celebration'

wan 'day'

chalẽ.m 'to celebrate'

we.la.no.n 'bedtime'

we.la. 'time'

no.n 'to sleep'

sanã.mlên 'playground'

sanã.m 'field'

lên 'to play'

wẽ.nmân 'engagement ring'

wẽ.n 'ring'

mân 'to be engaged'

?à.n?à.pnã.m 'bathtub'

?à.n 'bowl'

?à.pnã.m 'to bathe'

4.2.2.5 Object as Second Member. Finally, there are compounds of this type which preserve the underlying object as the second member. As it happens, the compounds rá·nkhă·y-kafe· 'coffee shop' and ro·nkèprót 'garage' which we used to illustrate the two derivations of Adverbial-Verb-Object compounds both have alternate forms without the verbs. rá·nkafe· also means 'coffee shop' and ro·nrót means 'garage'. Simple application of NCD 2 to the GNC-produced Adverbial-Verb-Object forms will yield the corresponding Adverbial-Object forms. Examples of Adverbial-Object compounds are:

bô·risàtyá·sup 'tobacco company'

bô·risàt 'company'

yá·sup 'tobacco'

rá·n?ahă·n 'restaurant'

rá·n 'shop'

?ahă·n 'food'

klôn·cunlathát 'microscope'

klôn 'pipe'

cunlathát 'microscopic object'

tû·sîaphâ· 'wardrobe, dresser'

tû· 'cabinet or similar container'

sîaphâ· 'clothing'

thănkhayà? 'garbage pail'

thăn 'bucket'

khayà? 'garbage'

banchi·nenfà·k 'bank account'

banchi· 'list, account'

nenfà·k 'deposit of money'

phâ·bay 'sailcloth, canvas'

phâ· 'cloth'

bay 'sail'

ro·ṇmahō·rasòp 'theatre'

ro·ṇ 'building'

mahō·rasòp 'entertainment'

hō·ka·nkhá· 'chamber of commerce'

hō· 'building erected for a special purpose'

ka·nkhá· 'trading'

hī·pburī· 'cigarette box'

hī·p 'box'

burī· 'cigarette'

4.2.3 Exocentric Compounds. In addition to the compounds

described above, there are also a few exocentric compounds whose heads do not appear. These are formed exactly like the compounds discussed so far, except that the head member, a Pro-noun, has been deleted by NCD 1. Some of these are listed below.

Subject Deleted, Object Preserved:

chō·mṇa·m 'beautiful girl'

chō·m 'appearance'

ṇa·m 'to be beautiful'

khwān?ō·n 'an easily-frightened person'

khwān 'spirits'

?ō·n 'to be tender'

phōmmá· 'bangs'

phōm 'hair'

má· 'horse'

să·mkhă· 'tripod'

să·m 'three'

khă· 'leg'

Object Deleted, Other Constituents Preserved:

cè·tcamnon 'aim, purpose'

cè·t 'mind'

camnon 'to desire'

kàpklê·m 'food eaten with alcoholic beverages'

kàp 'with'

klê·m 'snacks taken with drinks'

kàpkhâ·w 'food eaten with rice'

kàp 'with'

khâ·w 'rice'

ka·fà·k 'a kind of parasitic plant'

ka· 'crow'

fà·k 'to deposit'

lănchà·k 'inside story'

lăn 'back'

chà·k 'curtain'

Adverbial Deleted, Subject and Verb Preserved:

tawantòk 'West'

tawan 'sun'

tòk 'to drop'

tawan?ò·k 'East'

tawan 'sun'

?ò·k 'to emerge'

phèndinwăy 'earthquake'

phèndin 'earth'

wăy 'to tremble'

4.3 Review of the rules. At the close of the discussion, we will review the form of GNC and the Noun Compound Deletion transformations.

$$\text{GNC } \langle N \begin{cases} \text{a) } N \\ \text{b) } \langle \begin{matrix} (N) (V(\text{Prp})) \\ (V(\text{Prp})) (N) \end{matrix} \rangle_X \end{cases} \rangle_{NP}$$

$$\xrightarrow{1} \begin{matrix} 1 \\ 1 > 2 \end{matrix}$$

$$\text{NCD 1. } \langle (N \ V) \begin{bmatrix} +N \\ +\text{Pro} \end{bmatrix} (V \ N) \rangle_N$$

$$\xrightarrow{\quad} \begin{matrix} 1 & 2 & 3 \\ 1 & \emptyset & 3 \end{matrix}$$

Where 1 and 3 are neither both present nor both absent.

$$\text{NCD 2. } \langle N \ V(\text{Prp}) \ N \rangle_N$$

$$\xrightarrow{\quad} \begin{matrix} 1 & 2 & 3 & 4 \\ 1 & \emptyset & 3 & 4 \end{matrix}$$

$$\text{NCD 3. } \langle N \ (V) \ \text{Prp} \ N \rangle_N$$

$$\xrightarrow{\quad} \begin{matrix} 1 & 2 & 3 & 4 \\ 1 & 2 & \emptyset & 4 \end{matrix}$$

Obligatory if 2 is present; optional otherwise.

And, as an approximation of the principle dealing with the deletion of recompounded general forms:

$$\text{NCD 4. } \langle Z \langle \begin{matrix} \{XY\} \\ \{YX\} \end{matrix} \rangle_N \rangle_N$$

$$\xrightarrow{\quad} \begin{matrix} \langle \{XY\} \rangle_N Z \\ \{YX\} \end{matrix}$$

Delete X

Where X is a general term in an area of meaning in which Y is a specific term.

Footnotes

¹Lees, op. cit.

²A statement like "Noun modifiers in Thai always follow their head nouns, both in compounds and in noun phrases" is effectively two statements, of course. Lees, however, seems to have missed this generality since he places on his compound forming rule the burden of reversing the constituents in the underlying sentences. See Lees, op. cit., p. 174.

³Richard B. Noss, Thai Reference Grammar (Washington, D.C.: Foreign Service Institute, 1964), p. 23.

⁴Ibid.

⁵Personal communication.

⁶Udom Warotamasikkkhadit, Thai Syntax: An Outline (Bangkok: The College of Education Prasarnmitr, 1963), p. 45.

⁷Nisa Udomphol, Compound Words in Thai (Unpublished Master's Thesis, Chulalongkorn University, 1964).

⁸Ibid., p. 18.

⁹However, it is ungrammatical because of the presence of bay which is the classifier for cabinets, not keys. If the classifier dô·k is substituted the sentence could fail to be acceptable only because 'cabinet key' happens not to be commonly used. If the appropriate classifiers are allowed, this frame could never distinguish collocations from compounds.

¹⁰As we mentioned in footnote 9, the third frame is valueless if one takes seriously the generative capacity of compound formation in Thai. The second frame hinges on an interesting feature of Thai grammar which we cannot go into here. Suffice it to say that there are many sequences which will not fit into this frame which are not compounds either; hence it alone is not sufficient to distinguish noun phrases from compounds.

¹¹Udomphol, op. cit., p. 17.

¹²Ibid., p. 7-10.

¹³Pra Worawetpisit, Grammar (1954), p. 79. Quoted in Udomphol, op. cit., p. 7.

¹⁴Faculty of the Pre-University School, Grammar (The Pre-University School, 1956), p. 25. Quoted in Udomphol, op. cit., p. 9.

¹⁵Phya Upakitsilpasarn, Thai Grammar (1954), p. 3. Quoted in Udomphol, op. cit., p. 7-9.

¹⁶Reference Book of Thai Grammar for Secondary Education 4-5 (Ministry of Education, 1962), p. 202. Quoted in Udomphol, op. cit., p. 9.

¹⁷Campbell and Shaweevongse, op. cit., p. 153-155, 189-190, 255-257.

¹⁸Lanyon-Orgill, op. cit., p. 74-77.

¹⁹Anthony, et. al., op. cit., p. 358-359. Haas and Subhanka, op. cit., p. 25.

²⁰Noss, op. cit., p. 59-66.

²¹Mary R. Haas, Thai-English Student's Dictionary (Stanford, Calif.: Stanford University Press, 1964), p. xv-xvii.

²²Udomphol, op. cit.

²³Warotamasikkhadit, op. cit., p. 45-50.

²⁴Actually, the derivation is somewhat more complicated than this.

²⁵Udomphol, op. cit., p. 29.

²⁶Warotamasikkhadit, op. cit., p. 45.

²⁷Ibid. I will change Warotamasikkhadit's structural index in these rules slightly for greater clarity.

²⁸Ibid., p. 46.

²⁹Ibid., p. 49.

³⁰Haas, op. cit.

³¹Thai-Thai Dictionary (Bangkok: The Ministry of Education, 1950).

³²L.N. Morev (Comp.) and S. Semsampan (Ed.), Brief Thai-Russian Dictionary (Leningrad: Leningrad University Press, 1964).

³³Thai speakers may contest the classification of some of the examples in this chapter. Some are ambiguous and belong in more than one list. Some may be mis-classified. My only claim is 1) that the classification is valid and 2) that most of the examples in any list are correctly classified.

³⁴This compound involves coordinate compounding of two items which have already been coordinate-compounded.

³⁵Format is adapted from Rosenbaum, op. cit.

³⁶It is quite possible that the innermost sentence khon bin dūay sǎ·y is to be derived from the further embedding:

<khon cháy sǎmráp <ka·n <ka·n pen <<khon bin> S>NP>S>NP
 person use for activity activity be person fly
sǎ·y>_S
line

This type of derivation for instrumental phrases was suggested to me by Dr. James D. McCawley. For obvious reasons, we will not attempt to deal with this refinement in this discussion.

³⁷Frequently a noun position in a compound will be filled by a nominalized verb which has undergone T ka·n Del.

³⁸nák 'expert' is a bound form which must undergo GNC.

³⁹cha·w 'dweller in' is a bound form requiring application of GNC. Its presence in cha·wpramon is an enigma.

⁴⁰See footnote 38.

⁴¹For a definition for wuanom, the Thai-Thai Dictionary gives wua phan thí· hây nom má·k 'a kind of cow which gives much milk'. Thai-Thai Dictionary, op. cit.

⁴²Warotamasikkkhadit makes the same assumption implicitly with his rule that makes compounds out of sentences of the form: $N^1 V_t N^2$. See Warotamasikkkhadit, op. cit., p. 46 and pp. above.

⁴³Haas and Subhanka, op. cit., p. 551.

⁴⁴Haas and Subhanka, op. cit., pp. 549, 559.

⁴⁵Mr. Sirisamphan made the observation that these two sentences sounded like mistakes a European would make.

⁴⁶Cf. Warotamasikkkhadit, op. cit., p. 49-50 and discussion above.

⁴⁷Most of the other Thai government ministries are similarly named.

⁴⁸See footnote 38.

⁴⁹So startling was the first alternative that I at first assumed that he intended the second alternative to be a correction of the first, but he assured me that either definition was appropriate.

⁵⁰See footnote 39.

⁵¹A possible alternative to this and the tôn má·y 'tree' case would be to allow certain words, like tôn má·y and khâ·n·nāy to be entered as single lexical items with more than one member. This solution has a number of disadvantages. See chapter 5 for discussion of similar problems.

⁵²See footnote 41.

⁵³A very great many transitive verbs in Thai are ambivalent, as a glance at the meanings in the Thai-English Student's Dictionary will show. Warotamasikkkhadit (op. cit., p. 6, 7) sets up ambivalent verbs as a subset of Thai verbs separate from transitive verbs. For him there are obviously transitive verbs which are not ambivalent.

⁵⁴I actually did find a single example of a three-member compound; să·wkhonrāk 'a man's sweetheart' (să·w 'young woman', khon 'person', rāk 'to love'). There is nothing in our rules to prevent such compounds from being generated.

⁵⁵It seems likely that a more complete grammar of Thai would allow Pro-nouns to be more freely deleted. It may also be possible to collapse NCD 1 with T Pro Del. The presence of N in 3 is required by compound types to be discussed later.

⁵⁶Warotamasikkkhadit (op. cit., p. 48) derives phâ·măy 'silk' from phâ·tho·duay mǎy 'The cloth is woven with silk' instead of a sentence with tham 'to make'. It may be there are a number of verbs like this instead of just one.

⁵⁷The innermost sentence could undergo T Am Vb in the first cycle and thus khri·ankhrua could be derived like a Subject-Adverbial compound. For the sake of demonstration, we will assume that T Am Vb is not selected in the first cycle.

⁵⁸The constituent to·n 'Classifier for units of time' is deleted under the noun compound deletion principle.

⁵⁹The Ambivalent Verb Transformation could apply since khă·y is an ambivalent verb and the derivation would be simplified. For illustrative purposes we will assume that it does not.

Chapter 5

Learned Compounds

1.0 Introduction. Among noun compounds in Thai there is a small but interesting subset which differs from the others in several important ways. In the first place, the elements which make up these compounds are semi-assimilated elements of Indic origin, and are recognized as such by Thai speakers. Such compounds are pracha·sāksā 'public education' (Skt. prajā 'mankind', śikṣā 'training') and wisawako·n 'engineer' (Skt. viśva 'all', kara 'doer').¹ In the second place, the serial order of the elements in many of these is the reverse of what it would be if the compound were formed by the rules in chapter 4. The learned word phátta·kha·n 'large restaurant' is formed from the Indic elements phátta 'food' and ʔa·kha·n 'building' where the modifying element phátta precedes ʔa·kha·n, the head. By contrast, the native word for restaurant, rā·n-ʔa·hā·n (rā·n 'store', ʔa·hā·n 'food'), occurs in the normal order, head followed by modifier. Besides these factors, many of these learned compounds display atypical phonological phenomena. The word for 'animal' is sāt, and is so pronounced in isolation, or in compound with a native Thai noun, as in sātpa 'wild animal'. But when compounded with another Indic element in a learned compound it becomes sātwa-; sātwawittaya 'zoology' (witthaya 'science, -ology'). Similarly, the word

thun 'funds', when alone or compounded with certain other nouns has the form thun. But when compounded with sáp 'wealth', it becomes thunna (thunnasáp 'capital'). Other learned compounds show vestiges of Indic sandhi phenomena. The final vowel of the bound form phátta 'food' in the above-cited learned compound phátta·kha·n '(large) restaurant' fuses with the first vowel in ʔa·kha·n (with deletion of the glottal stop) to form the single long a·. There are a number of other examples, all involving short or long a.

2.0 Possible Solutions. It is clear that these items are special in a number of ways. What is less clear is how to account for the differences and still show how they fit into Thai grammar. One possible solution would be to enter them all in the lexicon as single units. This solution, however, runs into a number of serious objections. In a number of cases, the members of these learned compounds also occur independently with meanings similar to those they bear in compound. To enter the compounds as units would be to deny that the independent and compounded elements with the same meanings are the same item. Secondly, it would result in a multiplicity of repetition of the same items. We have already mentioned the word sàtawawitthaya· 'zoology'. But there are a number of academic disciplines that are named with compounds with witthaya·, and even more with sà·t, which also means 'science'. To enter each of these separately seems clearly to miss an obvious generality. Besides these objections, this solution

would build into the lexical entries the phonological alternations in forms like sàt and thun. While these alternations are irregular, there are a number of them that are irregular in the same way and the separate entry solution would preclude the possibility of exploiting these subregularities in the phonological rules. Finally, this type of solution would leave unexpressed the grammatical relationships between the members of these compounds, which are similar to those in the regular compounds.

Another possibility would be to allow the category symbol N to branch. A rule like $N \rightarrow (\text{Prefix}) N' (\text{Suffix})$ could be introduced into the grammar and the bound Indic elements could be specified as prefixes or suffixes depending on whether they follow or precede the head nouns of learned compounds. This solution allows us to avoid unnecessary repetition in the lexicon and would allow for expression of phonological subregularities. If the same form is allowed to be specified as a prefix or suffix and also a free form, there would be no necessity for double entries for items which appear either independently or in compound. But this solution would still fail to account for the grammatical relationships between members of the compounds, which provided much of the motivation for the derivation of compounds from sentences in the case of regular Thai compounds. In addition, this solution would force the retraction of the suggestion that categories like N and V are universal lexical category symbols which do not branch in the

grammar of any language.

Another possibility is to allow the compounds to be generated by the ordinary Thai rules for compound derivation, using native Thai forms, and then to allow substitution of the learned forms with concomitant reordering of these forms. This solution would seem to lack the defects of the other two, but has others of its own. In some cases, it is impossible for a Thai speaker to imagine a sentence which would account for the meaning and combination of learned forms in compound. For another thing, there seems to be a certain artificiality about such a solution in some instances. There are a great many compounds, especially of the learned type, of the form XY, for which the definition seems to be of the form X kław káp Y (or Y kław káp X) where kław káp means 'to deal with'. It appears that the concept of one thing dealing with, or having to do with another may be more basic than the subject-verb-object relationship captured by deriving compounds from underlying sentences. Perhaps the ultimate solution will be the recognition of an abstraction expressing the concept of X dealing with Y which may be expressed either by a sentence or by a normal Thai compound, or alternately, by a learned compound.

3.0 Informant Behaviour. Before we can adopt a tentative solution, it will be helpful to investigate what sort of knowledge a Thai speaker has on the basis of which he uses these compounds. The informant with whom I discussed these

matters was Kasian Chongsarit, a young man from Bangkok who has recently arrived in the United States to complete his final year of high school here before going on to college in an American university. The informant was first shown a few derivations for ordinary Thai compounds and then asked to supply underlying sentences for a few others of the native type. After doing this successfully, he was introduced to the learned type and asked to do two things. First, he was asked if he could distinguish the two parts which made up the word. He was able to do this without fail in every case. Secondly, he was asked to provide source sentences for this type of compound, using the two parts in the sentences if possible. The results of this task varied.

In many cases, he was simply unable to give a sentence at all. The word ?onka·n was identified as consisting of ?on 'body' and ka·n 'matter, affair'. But he was not able to suggest a sentence which reflected the meaning of these two parts. In others, a sentence was supplied which was difficult to evaluate. For kammasit 'ownership', Kasian suggested sitthi? thi· pen câwkhō·n "The situation which is being an owner". sitthi? 'situation' is an alternate form of sit, but the form kamma strictly speaking means 'deeds' and has no counterpart in the sentence. In some cases, he gave meaning for the constituent parts which it clearly does not have in the compound. In some of these cases, a form has a meaning

in compound which it does not have when used freely. For instance, kon in compound has the meaning 'device', as in koncàk 'engine'. Independently, however, it means 'trick'. The informant gave the meaning 'trick' for kon and was predictably unable to suggest an underlying sentence containing kon from which koncàk could be derived. In other cases, the informant seemed to be thinking of a word having nothing to do with the compound under discussion. For ?a·lay in wíthhaya·lay 'university', he gave the meaning "When someone that you like wants to depart and you don't want him to depart". This has nothing to do with its meaning in wíthhaya·lay, which is apparently close to the meaning of Sanskrit ālaya 'house, dwelling'.

It was interesting to note that the informant sometimes gave sentences as definitions for some learned compounds which would give meanings to the compound which differed substantially from the one given in the Thai-English Student's Dictionary. Almost always, these sentences followed the normal Thai pattern of head plus modifier, supplying a somewhat oblique verification of the analysis in chapter 4. One of these is the·walô·k 'the world of the gods' which consists of the·wa 'divinity, god' and lô·k 'world'. However, Kas'an gave the source as the·wada· thî· ráksā· lô·k "A divinity who protects the world" which interprets the head as the first member of the·walô·k instead of the second.

But in a large number of instances, the informant was

able to volunteer a definition which reflected an embedded sentence containing forms corresponding to the members of the compound and revealing the relevant grammatical relations. The learned word for 'artist', citrako·n (citra 'picture', ko·n 'doer') was defined as:

khon thi· wā·t rū·p
person who draw picture

"A person who draws pictures"

Cases of vowel sandhi were no more of a problem than the others. le·khă·nuka·n 'secretary' (le·khă· 'writing, ?anuka·n 'subordinate') received the definition:

khon thi· kīaw kāp ka·nkhian
person who deal with activity-write

"A person who deals with writing"

In other cases, the informant was not able to directly supply a definition, but accepted the one I suggested, sometimes with great enthusiasm. For phonlamian 'population' (phon(la) 'people', mian 'city, country'), I suggested:

khon thi· yù· nay mian
person who be-located in city, country

"People who are in the city (country)"

Kasian accepted this definition with the comment "That's good!"

4.0 Tentative Solution. Perhaps none of the solutions discussed in this chapter is ideal, but the most promising seems to be a combination of two of the above. It seems reasonable to assume in cases in which the informant was not able to suggest a sentence definition, and especially when he could not give a meaning to a compound member which elucidated its

4

function in the compound, that the putative compound is really a single lexical item. The difficulty here is that he always could correctly identify two items in these words. But it can be argued that this information is part of the Thai person's extralinguistic knowledge, analogous to the educated English speaker's knowledge that 'adventure' can be divided into 'ad' and 'venture' which can then be traced to Latin words. In neither case does this type of information contribute to efficiency in the description of the speaker's use of these words. Learned words for which satisfactory definitions could be given will then be derived as in chapter 4, using native forms. Late in the grammar, there will be a mechanism which allows the replacement of certain native forms with learned forms, followed by a transformation reversing the serial order where necessary. Thus, citrako·n 'artist' starts out as:

khon thi· wâ·t rû·p
 person who draw picture

"A person who draws pictures"

Regular grammatical processes convert this to khonrû·p. The substitution mechanism allows khon to be replaced by ko·n and rû·p by citra, giving *ko·ncitra. The reversal transformation makes the adjustment to citrako·n.

A similar solution takes the definitions given by the informant or dictionary less seriously. Instead of assuming that precisely the sentence given by the Thai speaker underlies the learned compound, it can be assumed that the sentence

is a paraphrase of the real underlying sentence necessitated by the presence in the real sentence of bound forms which cannot be used independently of the compound. The real underlying sentence, then, for a compound like citrako·n is not khon thi· wâ·t rû·p but *ko·n thi· wâ·t citra. The items ko·n and citra are specified in such a way that sentences containing them must undergo GNC in order to give a real Thai utterance. This solution would have the advantage that it can explain why ko·n 'doer' seems to have more specific semantic content than khon 'person'. It would also explain why Thai speakers will give the members of the compound in the definition-sentence, even if they are learned words, as long as they are free forms. A definition like khon thi· wâ·t rû·p is forced on the Thai speaker because he is not free to use ko·n and citra in a sentence. Nevertheless, it does reflect the grammatical relationships involved. If this solution is adopted, and it seems to be the best one available, the reversal rearrangement transformation would still be necessary, but the substitution mechanism would not.

However, sometimes the definitions given by the informant indicated grammatical relationships that are different from those indicated in the Thai government dictionary.² The word for 'capital city' râ·tchatha·ni· (râ·tcha 'king', tha·ni· 'city') was defined by my informant as:

mian thi· mahā·kasā·t yū·
city which king be-located
"A city where the king lives"

but by the dictionary as:

mian khō·n phrà·rà·teha·
city of king

"The city of the king"

and thus ultimately from:

mian thī· phrà·rà·teha· mī·
city which king have

"The city which the king has"

By the former definition, tha·nī· is the object of the preposition nay 'in'. In the other definition, it would be the direct object of the verb mī· 'have'. Similarly, mano·phā·p 'imagination' (mano· 'mind', phā·p 'picture') was defined by Kasian as:

phā·p thī· yū· nay khwa·mnikkhit
picture which be-located in essence-think

"A picture which is in the mind"

The dictionary, on the other hand, gives:

phā·p thī· nāk wā·t nay cītcaṃ
picture which be-thought be-drawn in mind

"A picture which is conceived in the mind"

According to the one definition, phā·p is ultimately the subject of yū· 'to be located' and according to the other, it is the ultimate object of nāk 'to think' and wā·t 'to draw'. But this state of affairs supports rather than destroys the analysis we propose. Since there are so many avenues of derivation which result in compounds, it is to be expected that many, if not most of them, will be ambiguous. This ambiguity can be accounted for by the above type of variations in definitions by native speakers of the language.

5.0 Examples. We will list the examples in the same order as in chapter 4, giving the structure-reflecting definition for each type, but no sample derivations since the derivations will be the same as those in chapter 4 except for the use of learned items and rearrangements where necessary. Each of the definitions given is attested either by an informant or in the Thai-Thai Dictionary. The proposed analysis for any item is valid only for such Thai speakers for whom the members of the compound and their interrelationships are known. Otherwise, the compound is functionally a unit for that speaker. By no means all the compound types listed in chapter 4 are found in the set of learned compounds. In particular, there will be no three-member compounds.

We include in chapter 5 a number of compounds which do not undergo rearrangement although they include learned members. We include them here since many of them are bound forms which appear only in learned compounds. We can consider these grammatically assimilated learned compounds.

5.1 Underlying Subject as Head.

5.1.1 Verb as Second Member.³

Examples which are not reversed:

Definition: samāybo·ra·n 'ancient times'

samāy thi· bo·ra·n
time which be-ancient

nāka·n 'cold sweat'

nā 'sweat'

ka·n 'to be black'

phanraya·nô·y 'minor or additional wife'

phanraya 'wife'

nô·y 'to be new'

hè·tsùtwisăy 'unavoidable circumstances'

hè·t 'cause, reason'

sùtwisăy 'to be beyond human power'

samăybo·ra·n 'ancient times'

samăy 'time'

bo·ra·n 'to be ancient'

sapha·să·man 'the House of Commons'

sapha 'council, assembly'

să·man 'to be common, ordinary'

ʔăksô·ntâm 'low consonant'

ʔăksô·n 'letter of the alphabet'

tâm 'to be low'

Examples which are reversed:

Definition: să·mansăksă 'elementary education'

ka·nsăksă thi să·man
education which be-ordinary

banpháburùt 'ancestors'

ban(phá) 'to be early, original'

burùt 'man'

bo·ra·nwátthù? 'relics'

bo·ra·n 'to be ancient'

wátthù? 'object'

lahù?thô·t 'light punishment'

lahù? 'to be light'

thô·t 'punishment'

mahăntaphay 'great danger'

mahănta 'to be great'

phay 'danger'

mahă·chon 'the public, the majority of the people'

mahă 'to be great'

chon 'people'

pàtchímmáway 'old age'

pàtchímmá 'to be final'

way 'period of life'

să·mansăksă 'elementary education'

să·man 'to be common, ordinary'

(ka·n)săksă 'education'

?ànúphanraya 'minor wife'

?ànú 'to be lesser, minor'

phanraya 'wife'

?ànúphâ·k 'particle'

?ànú 'to be lesser, minor'

phâ·k 'part'

?ùsăhà?kam 'industry'

?ùsăhà? 'to be diligent'

kam 'deeds'

5.1.2 Object as Second Member.

Examples which are not reversed:

Definition: khâ·râ·tchaka·n 'government official'

khon thî·râp râ·tchaka·n
person who work-for government

burùtpraysani 'mailman'

burùt 'man'

praysani 'mail'

kòtthammachâ·t 'laws of nature'

kòt 'rule, law'

thammachâ·t 'nature'

krabuanka·n 'procedure, administration'

krabuan 'procedure, procession'

ka·n 'matter'

khaná?kammaka·n 'committee'

khaná? 'group'

kammaka·n 'committee member'

khaná?ratthamontri· 'council of ministers'

khaná? 'group'

ratthamontri· 'minister in government'

khâ·lúan 'royal servant'

khâ· 'servant'

lúan 'government'

khâ·râ·tchaka·n 'government official'

khâ· 'servant'

râ·tchaka·n 'government'

prawátka·n 'history'

prawát 'history'

ka·n 'matter'

phrá?phu·m 'guardian spirit of one's land'

phrá? 'god'

phu·m 'earth'

?àthika·nwát 'abbot'

?àthika·n 'head, leader'

wát 'temple'

Examples which are reversed:

Definition: kitcaka·n 'activity'

ka·n thi· klaw káp thúra?
matter which deal with business

càkraya·n 'cycle' (the machine)

càkra 'wheel'

ya·n 'conveyance'

citrako·n 'artist'

citra 'painting'

ko·n 'doer'

kitcaka·n 'work, activity'

kitca 'work, activity'

ka·n 'activity'

khrittasā·tsanā· 'Christianity'

khrit(ta) 'Christ' (from European)

sā·tsanā· 'religion'

khunnasāp 'adjective'

khun(na) 'virtue'

sāp 'word'

phe·satchako·n 'pharmacist'

phe·sāt(cha) 'drug'

ko·n 'doer'

rátthamontri· 'government minister'

rát(tha) 'state, nation'

montri· 'minister'

rátthasapha· 'parliament'

rát(tha) 'state, nation'

sapha· 'council'

rô·ksínlapà? 'medicine, therapeutics'

rô·k 'disease'

sínlapà? 'art, craft'

yútthaka·n 'battle'

yút(tha) 'war'

ka·n 'activity'

An important subclass of Subject-Object learned compounds of which the members have been reversed are those which name academic disciplines with either sā·t or witthaya·, both of

which mean 'science' or '-ology' as head. Examples of this type are listed next.

Definition: sàtawáwitthaya 'zoology'
wícha thí klaw káp sàt
 science which deal with animal

cittawitthaya 'psychiatry'

cít(ta) 'mind'

witthaya 'science'

kasè·ttrasà·t 'agriculture as a science'

kasè·ttra 'cultivated field'

sà·t 'science'

khanittasà·t 'mathematics'

khanít(ta) 'mathematics'

sà·t 'science'

prawàtsà·t 'history as a discipline'

prawàt 'history'

sà·t 'science'

phráksasà·t 'botany'

phráksa 'plant'

sà·t 'science'

sānkhomwitthaya 'social science'

sānkhom 'society'

witthaya 'science'

sàtawáwitthaya 'zoology'

sàt(awá) 'animal'

witthaya 'science'

wē·tchasa·t 'medical science'

wē·tcha 'medicine'

sà·t 'science'

witthaya·sà·t 'science'

witthaya 'science'

sà·t 'science'

ʔaksō·nsà·t 'liberal arts'

ʔaksō·n 'letter of the alphabet'

sà·t 'science'

5.1.3 Adverbial as Second Member.

Examples which are not reversed:

Definition: phonlalô·k 'the peoples of the world'

khon thî·yù· nay lô·k
person who be-located in world

mano·phâ·p 'imagination'

mano· 'heart, mind'

phâ·p 'picture'

manútsayslô·k 'world of men'

manút(saya) 'man'

lô·k 'world'

phaythammachâ·t 'natural disaster'

phay 'danger'

thammachâ·t 'nature'

phonlalô·k 'the peoples of the world'

phonla 'people'

lô·k 'world'

phonlamian 'population'

phonla 'people'

mian 'city, country'

phonlarian 'civilian'

phonla 'person'

rian 'house'

thó·nphráʔro·n 'throne hall'

thó·n 'area'

phráʔro·n 'hall'

ʔamnâ·ttha·nka·nmian 'political power'
ʔamnâ·t 'power, influence'
tha·n 'associated with' (preposition)
ka·nmian 'politics'

Examples which are reversed:

Definition: râ·tchatha·ni· 'capital city'
mian thi· mahă·kasà·t yù·
city which king be-located

bannaphíphóp 'world of literature'
banna 'writings'
phíphóp 'world'

hima·lay 'Himalayas'
hima 'snow'
ʔa·lay 'place'

nakho·nsô·phe·ni· 'prostitute'
nakho·n 'city'
sô·phe·ni· 'prostitute'

râ·tchatha·ni· 'capital city'
râ·tcha 'king'
tha·ni· 'city'

5.1.4 Predicate Nominal as Second Member.

Examples which are not reversed:

Definition: rátthaba·nkla·n 'central government'
rátthaba·n thi· pen kla·n
government which be center

chonchâ·tthay 'Thai nationals'
chon 'people'
châ·tthay 'Thai race'

choncha·wʔankrít 'British people'
chon 'people'
cha·wʔankrít 'the British'

krabuanka·nhaydro·lisis 'hydrolysis'

krabuanka·n 'process'

haydro·lisis 'hydrolysis' (from English)

phontamrùat 'lowest rank of policeman'

phon 'member of a military or police force'

tamrùat 'policeman'

phonthahă·n 'private'

phon 'member of a military or police force'

thahă·n 'soldier'

phrá?borommara·chini· 'supreme queen'

phrá? 'ruler, god'

borommara·chini· 'supreme queen'

phrá?can 'the moon'

phrá? 'ruler, god'

can 'moon'

sămphanthamaytri· 'friendly relations'

sămphan(tha) 'relationship'

maytri· 'friendship'

?a·wútpi·n 'firearms'

?a·wút 'weapon'

pi·n 'gun'

Examples which are reversed:

Definition: mítracít 'friendliness'

cít thí· pen mít
mind which be friend

ca·rachon 'secret agent'

ca·ra 'spy'

chon 'person'

cltrakam 'painting' (art or act)

cltra 'painting'

kam 'deed'

maytri·cìt 'spirit of friendship'

maytri 'friendship'

cìt 'mind'

mìtracìt 'friendliness'

mìtra 'friend'

cìt 'mind'

Of the reversed Subject-Predicate Nominal examples, there is a special set of which the head member is phâ·p 'state, condition' or ka·n 'matter'. These nouns tend to add abstractness to the meaning of the other member. For various reasons, the usual sources for the underlying sentences for compounds was not productive for these. The informant, Kasian Chongsarit, was very reluctant to divide these compounds and give sentence-definitions including both members. When he did, it was clear that the sentence did not contribute to an understanding of the compound, and there was no consistency in his definitions. The Thai-Thai Dictionary is not likely to give definitions which include direct predicate-nominal relative clauses. Its definition for ʔìsaraphâ·p 'freedom' is:

<u>khwa·mpen</u>	<u>thay</u>	<u>kè·</u>	<u>tua</u>
quality-of-being	be-free	to	self

"State of being independent"

As we have seen in chapter 3, nominalizations of this type with khwa·m and ka·n are derived from underlying structures which involve sentences with predicate nominals at the deepest level. If we recognize khwa·m as the native Thai counterpart of phâ·p, and pen thay kè· tua as in some sense corresponding to ʔìsara, then we have the needed definition. Although the evidence is

sparse, it is difficult to see how the two members of these compounds could otherwise be related. It seems quite possible that the examples with ka·n are derived by the regular nominalization rule, except that 1) the underlying predicate nominal is a noun, not a sentence, and 2) the members are reversed by the reversal transformation for Indic compounds. It is also quite likely that at least some of these compounds are actually units for many Thai speakers. The examples are:

bùkkhalíkkaphâ·p 'personality'

bùkkhalíkka 'personality'

phâ·p 'state, condition'

khunnaphâ·p 'quality'

khun(na) 'virtue'

phâ·p 'state, condition'

mō·ranaphâ·p 'death'

mō·rana 'death'

phâ·p 'state, condition'

prawātka·n 'history'

prawāt 'account, history'

ka·n 'matter'

pháláka·n 'force, power'

phálá 'physical strength'

ka·n 'matter'

phě·nka·n 'plan, scheme'

phě·n 'plan'

ka·n 'matter'

sāmphanthaphâ·p 'relationship, state of alliance'

sāmphantha 'relationship'

phâ·p 'state, condition'

sāntiphā·p 'peacefulness'

sānti 'peace'

phā·p 'state, condition'

wonka·n 'circle, field, realm'

won 'circle'

ka·n 'matter'

ʔisaraphā·p 'freedom'

ʔisara 'to be free, freedom'

phā·p 'state, condition'

5.1.5 Exocentric Compounds.

Some learned compounds are exocentric in the sense that the head member of the compound does not actually appear in the compound. A few compounds of this type were mentioned in chapter 4, but proportionately more of them are found among learned compounds.

Definition: banna·rāk 'librarian'

<u>phū·</u>	<u>rāk</u>	<u>nansī·hōnsamūt</u>
person	protect	library-book

Examples of (Subject)-Object-Verb:

banna·rāk 'librarian'

banna· 'writings'

rāk 'to protect'

chonpratha·n 'irrigation'

chon 'water'

pratha·n 'to bestow'

nītibanyāt 'legislation'

nīti 'law'

banyāt 'to prescribe'

pracha·ba·n 'local government'

pracha· 'populace'

ba·n 'to tend'

rátthaba·n 'government'

rát(tha) 'nation, state'

ba·n 'to tend'

?onkharák 'body guard'

?on(kha) 'body'

rák 'to protect'

Example of (Subject)-Adverbial-Verb:

ya·mráksä·ka·n 'guard, sentry'

ya·m 'watches of the night' (Clf.)

ráksä·ka·n 'to be on guard duty'

Learned compounds in which underlying constituents other than the subject are the head are much rarer than subject-head compounds. Not only is the number of such compounds smaller, but there are fewer combinations of constituents and fewer compounds whose members are not reversed.

5.2 Underlying Object as Head.

5.2.1 Verb as Second Member.

Example which is reversed:

Definition: phalitphan 'product'

sín thí· phalit khén má·
thing which be-produced rise come

phalitphan 'product'

phalit 'to produce'

phan 'material'

5.2.2 Subject as Second Member.

Examples which are not reversed (all possessive):

Definition: mátimahă·chon 'public opinion'

khwa·mhén khô·n khonmù·mâ·k
opinion of large-group-people

mátimahă·chon 'public opinion'

mátì 'opinion'

mahă·chon 'populace'

prathê·tsarâ·t 'colony'

prathê·t(sa) 'country'

râ·t 'state, nation'

phônlakam 'retribution'

phôn(la) 'fruit, results'

kam 'deeds'

phu·míprathê·t 'topography'

phu·m(f) 'earth'

prathê·t 'country'

thamniapna·yók 'prime minister's residence'

thamniap 'residence of a high government official'

na·yók 'prime official'

?a·yúkla·nkhon 'middle age'

?a·yúkla·n 'middle age'

khon 'person'

?a·yúkhwa·m 'period of time after which prosecution of a crime
may no longer be initiated'

?a·yú 'age'

khwa·m 'case'

Examples which are reversed:

Definition: phrá?bà·t 'foot of a king or queen'

fá·thá·w khô·n phrá?câ·wphendin
sole-of-foot of ruler

ca·rakam 'espionage'

ca·ra 'spy'

kam 'deeds'

co·nrakam 'robbery'

co·n(ra) 'thief'

kam 'deeds'

manútsayatham 'humanity, humaneness'

manút(saya) 'man'

tham 'dharma'

phrá?bà·t 'foot of a king or queen'

phrá? 'god, ruler'

bà·t 'foot'

phrá?borommarû·p 'statue or picture of the king'

phrá? 'god, ruler'

borommarû·p 'splendid likeness'

phrá?thî·nân 'throne'

phrá? 'god, ruler'

thî·nân 'seat'

râ·tchathan 'official punishment'

râ·tcha 'king'

than 'punishment'

râ·tchathû·t 'royal envoy'

râ·tcha 'king'

thû·t 'ambassador'

sama·chîkphâ·p 'membership'

sama·chîk 'member'

phâ·p 'state, condition'

?a·yúkhăy 'life span'

?a·yú 'age'

khăy 'limit'

5.2.3 Adverbial as Second Member.

Example which is not reversed:

Definition: phátsadû?praysani· 'parcel post'

sínkhô·n thî· sôn tha·n praysani·
things which be-sent by mail

phátsadû?praysani· 'parcel post'

phátsadû? 'articles'

praysani· 'mail'

Examples which are reversed:

Definition: yútthaphan 'military equipment'

?a·wut thí· chày săm ràp tò·su.⁴
 weapon which be-used for fighting

mitsămphan 'friendly relations'

mít 'friend'

sămphan 'relationship'

pà·kkham 'spoken word'

pà·k 'mouth, beak'

kham 'word'

wê·tchaphan 'medical supplies'

wê·tcha 'medicine'

phan 'material'

yútthaphan 'military equipment'

yút(thā) 'war'

phan 'material'

?a·kà·tpraysani· 'airmail'

?a·kà·t 'air'

praysani· 'mail'

5.2.4 Exocentric Examples.

Examples of (Object)-Subject-Verbs:

Definition: kawī·nīphon 'poetry'

khamprāphan thí· kawī· tèn
 poetry which poet compose

kawī·nīphon 'poetry'

kawī· 'poet'

nīphon 'to write'

phrá?rá·tchadamrát 'royal speech'

phrá?rá·tcha 'king'

damrát 'to speak'

5.3 Underlying Adverbial as Head.5.3.1 Subject as Second Member.

Examples which are reversed:

Definition: sāṅkhommanūt 'human society'

manūt thī yū nay sāṅkhom
man which be-located in society

phūtthaka·n 'the time of Buddha'

phūt(tha) 'Buddha'

ka·n 'age'

sāṅkhommanūt 'human society'

sāṅkhom 'society'

manūt 'man'

the·walô·k 'heaven'

the·wa 'divinity'

lô·k 'world'

5.3.2 Object as Second Member.

Examples which are not reversed:

Definition: sathā·nmahō·rasòp 'theater'

thī· sāmrāp du·lên⁵
place for watch-for-amusement

sanā·mki·la· 'stadium, sports field'

sanā·m 'field'

ki·la· 'sport'

sathā·nmahō·rasòp 'theater'

sathā·n 'place'

mahō·rasòp 'entertainment'

sathā·nni·wittayú? 'radio station'

sathā·nni· 'station'

wittayú? 'radio'

Examples which are reversed:

Definition: witthaya·lay 'college'

sathă·n hây ka·nsăksă· chán sũ·n⁶
place give education level high

phátta·kha·n 'large restaurant'

phátta 'food'

?a·kha·n 'building'

witthaya·lay 'college'

witthaya· 'science, knowledge'

?a·lay 'place'

5.3.3 Exocentric Compound.

Example of (Adverbial)-Subject-Object:

pracha·thípatay 'democracy'

pracha· 'populace'

?athípatay 'sovereignty'

5.3.4 Co-ordinate Compounds.

Examples of co-ordinate compounds:

Definition: thúksũk 'happiness and sorrow'

thũk lé? sũk
sorrow and happiness

cítcay 'heart and soul'

cít 'mind'

cay 'heart'

phanraya·să·mí· 'married couple'

phanraya· 'wife'

să·mí· 'husband'

phayphibát 'danger, disaster'

phay 'danger'

phibát 'disaster'

pha·sí·?a·ko·n 'revenue'

pha·sí· 'tax'

?a·ko·n 'tax, revenue'

phrá?cā·w 'God, ruler'

phrá? 'god, ruler'

cā·w 'ruler, holy being'

rídu·ka·n 'season'

rídu· 'season'

ka·n 'age'

sāntisùk 'peace and happiness'

sānti 'peace'

sùk 'happiness'

sápaya·kō·n 'resources'

sáp(aya) 'wealth'

?a·kō·n 'revenue'

thē·pphacā·w 'a god'

thē·p(pha) 'divine being'

cā·w 'holy being, ruler'

thúksùk 'happiness and sorrow'

thúk 'sorrow'

sùk 'happiness'

5.3.5 Compounds Not Derived from Sentences.

As we found among the native Thai compounds, there are several learned compounds which are formed from noun phrases in which no sentence has been embedded. Again practically all examples occur with the number and noun in reverse order from the corresponding native compounds. The examples are:

banda·khru· 'all teachers'

banda· 'all of'

khru· 'teacher'

càtupàdcay 'four requisites for a Buddhist monk'

càtu 'four'

pàdcay 'requisite'

sàpphasàt 'all kinds of animals'

sàp(pha) 'all'

sàt 'animal'

sàpphakhun 'properties, qualities'

sàp(pha) 'all'

khun 'virtue'

?è·kkachon 'private individual'

?è·kka 'one'

chon 'person'

It will be noticed that sàpphasàt 'all kinds of animals' differs from the others in the list in that it has the meaning 'all kinds of animals' and not 'all animals'. This means that while the other examples have undergone the deletion of their regular classifiers, sàpphasàt has undergone deletion of a classifier meaning 'kind'. A similar example is phanmá·y 'kinds of plants' (phan 'kind' (clf.)-(tôn)má·y 'plant'). As a meaning for this word, Kasian gave me:

<u>tôn</u> <u>má·y</u>	<u>lan</u>	<u>chanít</u>
plant	several	kind

substituting the native classifier chanít for the learned phan. Clearly, phanmá·y and sàpphasàt are similar in derivation, except that sàpphasàt preserves the number word and deletes the classifier and phanmá·y represents the opposite process.

A number of these examples are exocentric ((Subject)-Object):

càtubà·t 'quadruped'

càtu 'four'

bà·t 'foot'

càturàt 'square'

càtu 'four'

ràt 'side'

sàpphaná·m 'pronoun'

sàp(pha) 'all'

ná·m 'noun'

5.3.6 Units. Finally, there are a number of words which, although they involve recurring parts, are judged to be units on the basis of informant behaviour and the lack of coherence between the meaning of the compound and the individual members.

Examples with càkra- 'wheel':

càkra-phát 'emperor'

càkra-phóp 'empire'

càkra-wa·n 'universe'

Examples with -ka·n 'matter':

kon-ka·n 'affair, business'

hè·t-ka·n 'event, circumstance'

ʔon-ka·n 'organization'

Examples with kam(ma) 'deeds' - kamma- preceding:

kamma-ka·n 'committee member'

kamma-ka·ròk 'accusative case'

kamma-phan 'heredity'

kamma-sít 'ownership'

kamma-wa·còk 'passive voice'

Examples with kam(ma) 'deeds' - kam following:

ka·yva-kam 'gymnastics'

sínlápa-kam 'art object'.

Footnotes

¹In discussing Thai vocabulary, Lanyon-Orgill notes: "...in the earlier period many Sanskrit and Pali learned words were introduced into Thai as a result of the Hindu influence in South-East Asia." Lanyon-Orgill, op. cit., p. 15.

²Thai-Thai Dictionary (Bangkok: Ministry of Education, 1950), passim.

³These differ from native compounds in that the verb is always an adjectival verb.

⁴Obviously, this involves the deeper embedding: khon tò·su· dūay a·wut "People fight with the weapons".

⁵The deeper embedding is khon du·lên nay thi· "People watch for amusement in the place".

⁶The deeper embedding is khon hây ka·nsàksă· chán sū·n nay sathă·n "People give high-level education at the place".

BIBLIOGRAPHY

Books

Thai Language.

- Allison, Gordon H. Modern Thai. Bangkok: Prachandra Printing Press, 1959.
- Anthony, Edward M. A Programmed Course in Reading Thai Syllables. Ann Arbor, Mich.: University of Michigan Press, 1962.
- Anthony, Edward M., Warotamasikkhadit, Udom, and French, Deborah P. Foundations of Thai, Book I. Pittsburgh: University of Pittsburgh, 1966.
- Campbell, Stuart, and Shaweevongse, Chuan. The Fundamentals of the Thai Language. (Second Edition). Bangkok: Thai Australia Company, 1956.
- Faculty of the Pre-University School. Grammar. Bangkok: The Pre-University School, 1955.
- Hass, Mary R. Thai-English Student's Dictionary. Stanford, Calif.: Stanford University Press, 1964.
- _____. Thai Reader. Washington, D.C.: American Council of Learned Societies, 1954.
- _____. The Thai System of Writing. Washington, D.C.: American Council of Learned Societies, 1956.
- Haas, Mary R. and Subhanka, Heng R. Spoken Thai. New York: Henry Holt and Co., 1945.
- Lanyon-Orgill, Peter A. An Introduction to the Thai (Siamese) Language for European Students. Victoria, B.C.: The Curlew Press, 1955.
- Morev, L.N. (Comp.) and Semsamphan, S. (ed.). Brief Thai-Russian Dictionary. Leningrad: Leningrad University Press, 1964.
- Noss, Richard B. Thai Reference Grammar. Washington, D.C.: Foreign Service Institute, 1964.
- Reference Book of Thai Grammar for Secondary Education 4-5. Bangkok: Ministry of Education, 1962.

Thai-Thai Dictionary. Bangkok: Ministry of Education, 1950.

Upakitsilpasarn, Phya. Thai Grammar. Bangkok, 1954.

Warotamasikkkhadit, Udom. Thai Syntax: An Outline. Bangkok: The College of Education Prasarnmitr, 1963.

Worwetpisit, Pra. Grammar. Bangkok, 1954.

Syntactic Theory.

Chomsky, Noam A. Aspects of the Theory of Syntax. Cambridge, Mass.: The M.I.T. Press, 1965.

_____. Syntactic Structures. The Hague, The Netherlands: Mouton and Co., 1957.

Katz, Jerrold J., and Postal, Paul M. An Integrated Theory of Linguistic Descriptions. Cambridge, Mass.: The M.I.T. Press, 1963.

Lakoff, George. On the Nature of Syntactic Irregularity. Cambridge, Mass.: The Computation Laboratory, Harvard University, Report NSF-16, 1965.

Lees, Robert B. The Grammar of English Nominalizations. Bloomington, Ind.: Indiana University Research Center in Anthropology, Folklore, and Linguistics, Revised Edition, 1963.

Rosenbaum, Peter S. The Grammar of English Predicate Complement Constructions. Cambridge, Mass.: The M.I.T. Press, 1967.

Articles and Unpublished Material

Thai Language.

Chaiyaratana, Chalao. "A Comparative Study of English and Thai Syntax." Unpublished Ph.D. dissertation, Indiana University, 1961.

Haas, Mary R. "Techniques of Intensifying in Thai," Word, II, 2 (1946), 127-130.

_____. "Types of Reduplication in Thai," Studies in Linguistics, I (1942), 1.4.1-1.4.6.

Haas, Mary R. "The Use of Numerical Classifiers in Thai," Language, XVIII (1942), 201-205.

Udomphol, Nisa. "Compound Words in Thai," Unpublished M.A. Thesis, Chulalongkorn University, 1962.

Syntactic Theory.

Anderson, Stephen R. "Concerning the Notion 'Base Component of a Transformational Grammar'." Cambridge, Mass., 1967. (Mimeographed.)

Fillmore, Charles J. "Concerning English Prepositions." Francis P. Dineen, S.J. (ed.), Report of the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies, Monograph Series on Languages and Linguistics No. 19: Washington, D.C.: Georgetown University Press, 1966.

Langendoen, D. Terence. "The Accessibility of Deep (Semantic) Structures," The Ohio State University, Working Papers in Linguistics, I (December, 1967), 119-127.

———. "On Selection, Projection, Meaning, and Semantic Content," The Ohio State University, Working Papers in Linguistics, I (December 1967), 100-108.

McCawley, James D. "Concerning the Base Component of a Transformational Grammar," Chicago, 1966. (Mimeographed)

———. "The Respective Downfalls of Deep Structure and Autonomous Syntax." Paper read before the 42nd Annual Meeting of the Linguistic Society of America, Chicago, Illinois, December 30, 1967.

———. "Where Do Noun Phrases Come From?" To appear in Jacobs, Roderick, and Rosenbaum, Peter S. (eds.), Readings in English Transformational Grammar.

Ross, John Robert. "Adjectives as Noun Phrases," Cambridge, Mass., 1966. (Mimeographed.)